

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2001, 20:37:17 ; Search time 1583.92 Seconds
(without alignments)
3735.975 Million cell updates/sec

Title: US-09-691-763B-4
Perfect score: 626
Sequence: 1 agcgcttttctgagggc.....ggaagggaaggatcactt 626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*
10: gb_est10.*
11: gb_est11.*
12: gb_est12.*
13: gb_est13.*
14: gb_est14.*
15: gb_est15.*
16: gb_est16.*
17: gb_est17.*
18: gb_est18.*
19: gb_est19.*
20: gb_est20.*
21: gb_est21.*
22: gb_est22.*
23: gb_est23.*
24: gb_est24.*
25: gb_est25.*
26: gb_est26.*
27: gb_est27.*
28: gb_est28.*
29: gb_est29.*
30: gb_est30.*
31: gb_est31.*
32: gb_est32.*
33: gb_est33.*
34: gb_est34.*
35: gb_est35.*
36: gb_est36.*
37: gb_est37.*
38: gb_est38.*
39: gb_est39.*
40: gb_est40.*
41: gb_est41.*
42: gb_est42.*
43: gb_est43.*
44: gb_est44.*
45: gb_est45.*
46: gb_est46.*
47: gb_est47.*
48: gb_est48.*
49: gb_est49.*
50: gb_est50.*
51: gb_est51.*
52: gb_est52.*
53: gb_est53.*
54: gb_est54.*
55: gb_est55.*
56: gb_est56.*
57: gb_est57.*
58: gb_est58.*
59: gb_est59.*
60: gb_est60.*
61: gb_est61.*
62: gb_est62.*
63: gb_est63.*
64: gb_est64.*
65: gb_est65.*
66: gb_est66.*
67: gb_est67.*
68: gb_est68.*
69: gb_est69.*
70: gb_est70.*
71: gb_est71.*
72: gb_est72.*
73: gb_est73.*
74: gb_est74.*
75: gb_est75.*
76: gb_est76.*
77: gb_est77.*
78: gb_est78.*
79: gb_est79.*
80: gb_est80.*
81: gb_est81.*
82: gb_est82.*
83: gb_est83.*
84: gb_est84.*
85: gb_est85.*
86: gb_est86.*
87: gb_est87.*
88: gb_est88.*
89: gb_est89.*
90: gb_est90.*
91: gb_est91.*
92: gb_est92.*
93: gb_est93.*
94: gb_est94.*
95: gb_est95.*
96: gb_est96.*
97: gb_est97.*
98: gb_est98.*
99: gb_est99.*
100: gb_est100.*
101: gb_est101.*
102: gb_est102.*
103: gb_est103.*
104: gb_est104.*
105: gb_est105.*
106: gb_est106.*
107: gb_est107.*
108: gb_est108.*
109: gb_est109.*
110: gb_est110.*
111: gb_est111.*
112: gb_est112.*
113: gb_est113.*
114: gb_est114.*
115: gb_est115.*
116: gb_est116.*
117: gb_est117.*
118: gb_est118.*
119: gb_est119.*
120: gb_est120.*
121: gb_est121.*
122: gb_est122.*
123: gb_est123.*
124: gb_est124.*
125: gb_est125.*
126: gb_est126.*
127: gb_est127.*
128: gb_est128.*
129: gb_est129.*
130: gb_est130.*
131: gb_est131.*
132: gb_est132.*
133: gb_est133.*
134: gb_est134.*
135: gb_est135.*
136: gb_est136.*
137: gb_est137.*
138: gb_est138.*
139: gb_est139.*
140: gb_est140.*
141: gb_est141.*
142: gb_est142.*
143: gb_est143.*
144: gb_est144.*
145: gb_est145.*
146: gb_est146.*
147: gb_est147.*
148: gb_est148.*
149: gb_est149.*
150: gb_est150.*
151: gb_est151.*
152: gb_est152.*
153: gb_est153.*
154: gb_est154.*
155: gb_est155.*
156: gb_est156.*
157: gb_est157.*
158: gb_est158.*
159: gb_est159.*
160: gb_est160.*
161: gb_est161.*
162: gb_est162.*
163: gb_est163.*
164: gb_est164.*
165: gb_est165.*
166: gb_est166.*
167: gb_est167.*
168: gb_est168.*
169: gb_est169.*
170: gb_est170.*
171: gb_est171.*
172: gb_est172.*
173: gb_est173.*
174: gb_est174.*
175: gb_est175.*
176: gb_est176.*
177: gb_est177.*
178: gb_est178.*
179: gb_est179.*
180: gb_est180.*
181: gb_est181.*
182: gb_est182.*
183: gb_est183.*
184: gb_est184.*
185: gb_est185.*
186: gb_est186.*
187: gb_est187.*
188: gb_est188.*
189: gb_est189.*
190: gb_est190.*
191: gb_est191.*
192: gb_est192.*
193: gb_est193.*
194: gb_est194.*
195: gb_est195.*
196: gb_est196.*
197: gb_est197.*
198: gb_est198.*
199: gb_est199.*
200: gb_est200.*
201: gb_est201.*
202: gb_est202.*
203: gb_est203.*
204: gb_est204.*
205: gb_est205.*
206: gb_est206.*
207: gb_est207.*
208: gb_est208.*
209: gb_est209.*
210: gb_est210.*
211: gb_est211.*
212: gb_est212.*
213: gb_est213.*
214: gb_est214.*
215: gb_est215.*
216: gb_est216.*
217: gb_est217.*
218: gb_est218.*
219: gb_est219.*
220: gb_est220.*
221: gb_est221.*
222: gb_est222.*
223: gb_est223.*
224: gb_est224.*
225: gb_est225.*
226: gb_est226.*
227: gb_est227.*
228: gb_est228.*
229: gb_est229.*
230: gb_est230.*
231: gb_est231.*
232: gb_est232.*
233: gb_est233.*
234: gb_est234.*
235: gb_est235.*
236: gb_est236.*
237: gb_est237.*
238: gb_est238.*
239: gb_est239.*
240: gb_est240.*
241: gb_est241.*
242: gb_est242.*
243: gb_est243.*
244: gb_est244.*
245: gb_est245.*
246: gb_est246.*
247: gb_est247.*
248: gb_est248.*
249: gb_est249.*
250: gb_est250.*
251: gb_est251.*
252: gb_est252.*
253: gb_est253.*
254: gb_est254.*
255: gb_est255.*
256: gb_est256.*
257: gb_est257.*
258: gb_est258.*
259: gb_est259.*
260: gb_est260.*
261: gb_est261.*
262: gb_est262.*
263: gb_est263.*
264: gb_est264.*
265: gb_est265.*
266: gb_est266.*
267: gb_est267.*
268: gb_est268.*
269: gb_est269.*
270: gb_est270.*
271: gb_est271.*
272: gb_est272.*
273: gb_est273.*
274: gb_est274.*
275: gb_est275.*
276: gb_est276.*
277: gb_est277.*
278: gb_est278.*
279: gb_est279.*
280: gb_est280.*
281: gb_est281.*
282: gb_est282.*
283: gb_est283.*
284: gb_est284.*
285: gb_est285.*
286: gb_est286.*
287: gb_est287.*
288: gb_est288.*
289: gb_est289.*
290: gb_est290.*
291: gb_est291.*
292: gb_est292.*
293: gb_est293.*
294: gb_est294.*
295: gb_est295.*
296: gb_est296.*
297: gb_est297.*
298: gb_est298.*
299: gb_est299.*
300: gb_est300.*
301: gb_est301.*
302: gb_est302.*
303: gb_est303.*
304: gb_est304.*
305: gb_est305.*
306: gb_est306.*
307: gb_est307.*
308: gb_est308.*
309: gb_est309.*
310: gb_est310.*
311: gb_est311.*
312: gb_est312.*
313: gb_est313.*
314: gb_est314.*
315: gb_est315.*
316: gb_est316.*
317: gb_est317.*
318: gb_est318.*
319: gb_est319.*
320: gb_est320.*
321: gb_est321.*
322: gb_est322.*
323: gb_est323.*
324: gb_est324.*
325: gb_est325.*
326: gb_est326.*
327: gb_est327.*
328: gb_est328.*
329: gb_est329.*
330: gb_est330.*
331: gb_est331.*
332: gb_est332.*
333: gb_est333.*
334: gb_est334.*
335: gb_est335.*
336: gb_est336.*
337: gb_est337.*
338: gb_est338.*
339: gb_est339.*
340: gb_est340.*
341: gb_est341.*
342: gb_est342.*
343: gb_est343.*
344: gb_est344.*
345: gb_est345.*
346: gb_est346.*
347: gb_est347.*
348: gb_est348.*
349: gb_est349.*
350: gb_est350.*
351: gb_est351.*
352: gb_est352.*
353: gb_est353.*
354: gb_est354.*
355: gb_est355.*
356: gb_est356.*
357: gb_est357.*
358: gb_est358.*
359: gb_est359.*
360: gb_est360.*
361: gb_est361.*
362: gb_est362.*
363: gb_est363.*
364: gb_est364.*
365: gb_est365.*
366: gb_est366.*
367: gb_est367.*
368: gb_est368.*
369: gb_est369.*
370: gb_est370.*
371: gb_est371.*
372: gb_est372.*
373: gb_est373.*
374: gb_est374.*
375: gb_est375.*
376: gb_est376.*
377: gb_est377.*
378: gb_est378.*
379: gb_est379.*
380: gb_est380.*
381: gb_est381.*
382: gb_est382.*
383: gb_est383.*
384: gb_est384.*
385: gb_est385.*
386: gb_est386.*
387: gb_est387.*
388: gb_est388.*
389: gb_est389.*
390: gb_est390.*
391: gb_est391.*
392: gb_est392.*
393: gb_est393.*
394: gb_est394.*
395: gb_est395.*
396: gb_est396.*
397: gb_est397.*
398: gb_est398.*
399: gb_est399.*
400: gb_est400.*
401: gb_est401.*
402: gb_est402.*
403: gb_est403.*
404: gb_est404.*
405: gb_est405.*
406: gb_est406.*
407: gb_est407.*
408: gb_est408.*
409: gb_est409.*
410: gb_est410.*
411: gb_est411.*
412: gb_est412.*
413: gb_est413.*
414: gb_est414.*
415: gb_est415.*
416: gb_est416.*
417: gb_est417.*
418: gb_est418.*
419: gb_est419.*
420: gb_est420.*
421: gb_est421.*
422: gb_est422.*
423: gb_est423.*
424: gb_est424.*
425: gb_est425.*
426: gb_est426.*
427: gb_est427.*
428: gb_est428.*
429: gb_est429.*
430: gb_est430.*
431: gb_est431.*
432: gb_est432.*
433: gb_est433.*
434: gb_est434.*
435: gb_est435.*
436: gb_est436.*
437: gb_est437.*
438: gb_est438.*
439: gb_est439.*
440: gb_est440.*
441: gb_est441.*
442: gb_est442.*
443: gb_est443.*
444: gb_est444.*
445: gb_est445.*
446: gb_est446.*
447: gb_est447.*
448: gb_est448.*
449: gb_est449.*
450: gb_est450.*
451: gb_est451.*
452: gb_est452.*
453: gb_est453.*
454: gb_est454.*
455: gb_est455.*
456: gb_est456.*
457: gb_est457.*
458: gb_est458.*
459: gb_est459.*
460: gb_est460.*
461: gb_est461.*
462: gb_est462.*
463: gb_est463.*
464: gb_est464.*
465: gb_est465.*
466: gb_est466.*
467: gb_est467.*
468: gb_est468.*
469: gb_est469.*
470: gb_est470.*
471: gb_est471.*
472: gb_est472.*
473: gb_est473.*
474: gb_est474.*
475: gb_est475.*
476: gb_est476.*
477: gb_est477.*
478: gb_est478.*
479: gb_est479.*
480: gb_est480.*
481: gb_est481.*
482: gb_est482.*
483: gb_est483.*
484: gb_est484.*
485: gb_est485.*
486: gb_est486.*
487: gb_est487.*
488: gb_est488.*
489: gb_est489.*
490: gb_est490.*
491: gb_est491.*
492: gb_est492.*
493: gb_est493.*
494: gb_est494.*
495: gb_est495.*
496: gb_est496.*
497: gb_est497.*
498: gb_est498.*
499: gb_est499.*
500: gb_est500.*
501: gb_est501.*
502: gb_est502.*
503: gb_est503.*
504: gb_est504.*
505: gb_est505.*
506: gb_est506.*
507: gb_est507.*
508: gb_est508.*
509: gb_est509.*
510: gb_est510.*
511: gb_est511.*
512: gb_est512.*
513: gb_est513.*
514: gb_est514.*
515: gb_est515.*
516: gb_est516.*
517: gb_est517.*
518: gb_est518.*
519: gb_est519.*
520: gb_est520.*
521: gb_est521.*
522: gb_est522.*
523: gb_est523.*
524: gb_est524.*
525: gb_est525.*
526: gb_est526.*
527: gb_est527.*
528: gb_est528.*
529: gb_est529.*
530: gb_est530.*
531: gb_est531.*
532: gb_est532.*
533: gb_est533.*
534: gb_est534.*
535: gb_est535.*
536: gb_est536.*
537: gb_est537.*
538: gb_est538.*
539: gb_est539.*
540: gb_est540.*
541: gb_est541.*
542: gb_est542.*
543: gb_est543.*
544: gb_est544.*
545: gb_est545.*
546: gb_est546.*
547: gb_est547.*
548: gb_est548.*
549: gb_est549.*
550: gb_est550.*
551: gb_est551.*
552: gb_est552.*
553: gb_est553.*
554: gb_est554.*
555: gb_est555.*
556: gb_est556.*
557: gb_est557.*
558: gb_est558.*
559: gb_est559.*
560: gb_est560.*
561: gb_est561.*
562: gb_est562.*
563: gb_est563.*
564: gb_est564.*
565: gb_est565.*
566: gb_est566.*
567: gb_est567.*
568: gb_est568.*
569: gb_est569.*
570: gb_est570.*
571: gb_est571.*
572: gb_est572.*
573: gb_est573.*
574: gb_est574.*
575: gb_est575.*
576: gb_est576.*
577: gb_est577.*
578: gb_est578.*
579: gb_est579.*
580: gb_est580.*
581: gb_est581.*
582: gb_est582.*
583: gb_est583.*
584: gb_est584.*
585: gb_est585.*
586: gb_est586.*
587: gb_est587.*
588: gb_est588.*
589: gb_est589.*
590: gb_est590.*
591: gb_est591.*
592: gb_est592.*
593: gb_est593.*
594: gb_est594.*
595: gb_est595.*
596: gb_est596.*
597: gb_est597.*
598: gb_est598.*
599: gb_est599.*
600: gb_est600.*
601: gb_est601.*
602: gb_est602.*
603: gb_est603.*
604: gb_est604.*
605: gb_est605.*
606: gb_est606.*
607: gb_est607.*
608: gb_est608.*
609: gb_est609.*
610: gb_est610.*
611: gb_est611.*
612: gb_est612.*
613: gb_est613.*
614: gb_est614.*
615: gb_est615.*
616: gb_est616.*
617: gb_est617.*
618: gb_est618.*
619: gb_est619.*
620: gb_est620.*
621: gb_est621.*
622: gb_est622.*
623: gb_est623.*
624: gb_est624.*
625: gb_est625.*
626: gb_est626.*
627: gb_est627.*
628: gb_est628.*
629: gb_est629.*
630: gb_est630.*
631: gb_est631.*
632: gb_est632.*
633: gb_est633.*
634: gb_est634.*
635: gb_est635.*
636: gb_est636.*
637: gb_est637.*
638: gb_est638.*
639: gb_est639.*
640: gb_est640.*
641: gb_est641.*
642: gb_est642.*
643: gb_est643.*
644: gb_est644.*
645: gb_est645.*
646: gb_est646.*
647: gb_est647.*
648: gb_est648.*
649: gb_est649.*
650: gb_est650.*
651: gb_est651.*
652: gb_est652.*
653: gb_est653.*
654: gb_est654.*
655: gb_est655.*
656: gb_est656.*
657: gb_est657.*
658: gb_est658.*
659: gb_est659.*
660: gb_est660.*
661: gb_est661.*
662: gb_est662.*
663: gb_est663.*
664: gb_est664.*
665: gb_est665.*
666: gb_est666.*
667: gb_est667.*
668: gb_est668.*
669: gb_est669.*
670: gb_est670.*
671: gb_est671.*
672: gb_est672.*
673: gb_est673.*
674: gb_est674.*
675: gb_est675.*
676: gb_est676.*
677: gb_est677.*
678: gb_est678.*
679: gb_est679.*
680: gb_est680.*
681: gb_est681.*
682: gb_est682.*
683: gb_est683.*
684: gb_est684.*
685: gb_est685.*
686: gb_est686.*
687: gb_est687.*
688: gb_est688.*
689: gb_est689.*
690: gb_est690.*
691: gb_est691.*
692: gb_est692.*
693: gb_est693.*
694: gb_est694.*
695: gb_est695.*
696: gb_est696.*
697: gb_est697.*
698: gb_est698.*
699: gb_est699.*
700: gb_est700.*
701: gb_est701.*
702: gb_est702.*
703: gb_est703.*
704: gb_est704.*
705: gb_est705.*
706: gb_est706.*
707: gb_est707.*
708: gb_est708.*
709: gb_est709.*
710: gb_est710.*
711: gb_est711.*
712: gb_est712.*
713: gb_est713.*
714: gb_est714.*
715: gb_est715.*
716: gb_est716.*
717: gb_est717.*
718: gb_est718.*
719: gb_est719.*
720: gb_est720.*
721: gb_est721.*
722: gb_est722.*
723: gb_est723.*
724: gb_est724.*
725: gb_est725.*
726: gb_est726.*
727: gb_est727.*
728: gb_est728.*
729: gb_est729.*
730: gb_est730.*
731: gb_est731.*
732: gb_est732.*
733: gb_est733.*
734: gb_est734.*
735: gb_est735.*
736: gb_est736.*
737: gb_est737.*
738: gb_est738.*
739: gb_est739.*
740: gb_est740.*
741: gb_est741.*
742: gb_est742.*
743: gb_est743.*
744: gb_est744.*
745: gb_est745.*
746: gb_est746.*
747: gb_est747.*
748: gb_est748.*
749: gb_est749.*
750: gb_est750.*
751: gb_est751.*
752: gb_est752.*
753: gb_est753.*
754: gb_est754.*
755: gb_est755.*
756: gb_est756.*
757: gb_est757.*
758: gb_est758.*
759: gb_est759.*
760: gb_est760.*
761: gb_est761.*
762: gb_est762.*
763: gb_est763.*
764: gb_est764.*
765: gb_est765.*
766: gb_est766.*
767: gb_est767.*
768: gb_est768.*
769: gb_est769.*
770: gb_est770.*
771: gb_est771.*
772: gb_est772.*
773: gb_est773.*
774: gb_est774.*
775: gb_est775.*
776: gb_est776.*
777: gb_est777.*
778: gb_est778.*
779: gb_est779.*
780: gb_est780.*
781: gb_est781.*
782: gb_est782.*
783: gb_est783.*
784: gb_est784.*
785: gb_est785.*
786: gb_est786.*
787: gb_est787.*
788: gb_est788.*
789: gb_est789.*
790: gb_est790.*
791: gb_est791.*
792: gb_est792.*
793: gb_est793.*
794: gb_est794.*
795: gb_est795.*
796: gb_est796.*
797: gb_est797.*
798: gb_est798.*
799: gb_est799.*
800: gb_est800.*
801: gb_est801.*
802: gb_est802.*
803: gb_est803.*
804: gb_est804.*
805: gb_est805.*
806: gb_est806.*
807: gb_est807.*
808: gb_est808.*
809: gb_est809.*
810: gb_est810.*
811: gb_est811.*
812: gb_est812.*
813: gb_est813.*
814: gb_est814.*
815: gb_est815.*
816: gb_est816.*
817: gb_est817.*
818: gb_est818.*
819: gb_est819.*
820: gb_est820.*
821: gb_est821.*
822: gb_est822.*
823: gb_est823.*
824: gb_est824.*
825: gb_est825.*
826: gb_est826.*
827: gb_est827.*
828: gb_est828.*
829: gb_est829.*
830: gb_est830.*
831: gb_est831.*
832: gb_est832.*
833: gb_est833.*
834: gb_est834.*
835: gb_est835.*
836: gb_est836.*
837: gb_est837.*
838: gb_est838.*
839: gb_est839.*
840: gb_est840.*
841: gb_est841.*
842: gb_est842.*
843: gb_est843.*
844: gb_est844.*
845: gb_est845.*
846: gb_est846.*
847: gb_est847.*
848: gb_est848.*
849: gb_est849.*
850: gb_est850.*
851: gb_est851.*
852: gb_est852.*
853: gb_est853.*
854: gb_est854.*
855: gb_est855.*
856: gb_est856.*
857: gb_est857.*
858: gb_est858.*
859: gb_est859.*
860: gb_est860.*
861: gb_est861.*
862: gb_est862.*
863: gb_est863.*
864: gb_est864.*
865: gb_est865.*
866: gb_est866.*
867: gb_est867.*
868: gb_est868.*
869: gb_est869.*
870: gb_est870.*
871: gb_est871.*
872: gb_est872.*
873: gb_est873.*
874: gb_est874.*
875: gb_est875.*
876: gb_est876.*
877: gb_est877.*
878: gb_est878.*
879: gb_est879.*
880: gb_est880.*
881: gb_est881.*
882: gb_est882.*
883: gb_est883.*
884: gb_est884.*
885: gb_est885.*
886: gb_est886.*
887: gb_est887.*
888: gb_est888.*
889: gb_est889.*
890: gb_est890.*
891: gb_est891.*
892: gb_est892.*
893: gb_est893.*
894: gb_est894.*
895: gb_est895.*
896: gb_est896.*
897: gb_est897.*
898: gb_est898.*
899: gb_est899.*
900: gb_est900.*
901: gb_est901.*
902: gb_est902.*
903: gb_est903.*
904: gb_est904.*
905: gb_est905.*
906: gb_est906.*
907: gb_est907.*
908: gb_est908.*
909: gb_est909.*
910: gb_est910.*
911: gb_est911.*
912: gb_est912.*
913: gb_est913

REFERENCE 1 (bases 1 to 888)
 TITLE NIH-MGC http://mgc.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M9709 row: e column: 23
 High quality sequence stop: 659.

FEATURES
 source
 Location/Qualifiers
 1..888

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3904126"
 /clone_lib="NIH_MGC_70"
 /tissue_type="epithelioid carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: Pancreas; Vector: PCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 Kb. Library constructed by Life Technologies."

BASE COUNT 217 a 236 c 281 g 154 t
 ORIGIN

Query Match 41.4%; Score 259; DB 141; Length 888;
 Best Local Similarity 100.0%; Pred. No. 1.6e-42;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 172 gacgcacatcctggatgcgctggagacacctgacccgcgaggagctcaagaagttcaagctg 231
 Db 1 GAGCGCATCTCGATGCGCTGGAGAACCTGACCGCGGAGGAGCTCAAGAAGTTCAAGCTG 60
 QY 232 aagctgctcgtgctgcgctgcgagggctacggcgctaccccgcgggcgctgctg 291
 Db 61 AAGCTCTGCTGCTGCTGCTGCGAGGGCTACGGCGGCATCCCGCGGGCGCGCTGCTG 120
 QY 292 tcatgagcccttggaacctaccgacacagctggctcagctctcctacctggagacctcagcg 351
 Db 121 TCCATGGAGCGCTTGGACCTCACCAGCAAGCTGGTCACTTCTTACCTGGAGACCTACGGC 180
 QY 352 gccagctcaccgctaacgtgctgcgacatggcctgcaggagatggccggcgagctg 411
 Db 181 GCGGAGCTCACCCTTAACTGCTGCGGACATGGGCTTGCAGGAGATGGCGGGCAGCTG 240
 QY 412 caggcgccacgacccagg 430
 Db 241 CAGCGCGCCACGACACAGG 259

RESULT 7
 BE908204
 LOCUS A1148558 678 bp mRNA EST 28-OCT-1998
 DEFINITION qc68g10.x1 Soares.Placenta.8to9weeks.2NBP8to9w Homo sapiens cDNA clone IMAGE:1714818 3' similar to TR:015553 015553 PYRIN. ;, mRNA sequence.

ACCESSION A1148558
 VERSION A1148558.1 GI:3677027
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 678)
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE
 JOURNAL
 COMMENT

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 804 Std Error: 0.00
 Seq primer: -40m13 fwd. EF from Amersham
 High quality sequence stop: 402.

FEATURES
 source

Location/Qualifiers
 1..678
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:1714818"
 /clone_lib="Soares.Placenta.8to9weeks.2NBP8to9w"
 /dev_stage="two placentae; one from 8 weeks and another from 9 weeks post conception"
 /lab_host="DH10B (ampicillin resistant)"
 /note="Organ: placenta; Vector: p7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGAGCGCGCGGATTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Facima Bonaldo."

BASE COUNT 114 a 202 c 217 g 145 t
 ORIGIN

Query Match 41.3%; Score 258.4; DB 16; Length 678;
 Best Local Similarity 99.6%; Pred. No. 2.1e-42;
 Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 172 gacgcacatcctggatgcgctggagacacctgacccgcgaggagctcaagaagttcaagctg 231
 Db 678 GACGCCATCTCGATGCGCTGGAGAACCTGCGCGGAGGAGCTCAAGAAGTTCAAGCTG 619
 QY 232 aagctgctcgtgctgcgctgcgagggctacggcgctaccccgcgggcgctgctg 291
 Db 618 AAGCTCTGCTGCTGCTGCTGCGAGGGCTACGGCGGCATCCCGCGGGCGCGCTGCTG 559
 QY 292 tcatgagcccttggaacctaccgacacagctggctcagctctcctacctggagacctcagcg 351
 Db 558 TCCATGGAGCGCTTGGACCTCACCAGCAAGCTGGTCACTTCTTACCTGGAGACCTACGGC 499
 QY 352 gccagctcaccgctaacgtgctgcgacatggcctgcaggagatggccggcgagctg 411
 Db 498 GCGGAGCTCACCCTTAACTGCTGCGGACATGGGCTTGCAGGAGATGGCGGGCAGCTG 439
 QY 412 caggcgccacgacccagg 431
 Db 438 CAGCGCGCCACGACACAGG 419

RESULT 8
 BE908204
 LOCUS 601500461F1 NIH_MGC_70 Homo sapiens cDNA clone IMAGE:3902319 5', mRNA sequence.
 ACCESSION BE908204
 VERSION BE908204.1 GI:10402543
 KEYWORDS EST.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 675)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

Fri Aug 31 11:08:29 2001

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1AM9704 row: j column: 16
High quality sequence stop: 660.
Location/Qualifiers
1. 675
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3902319"
/tissue_type="epithelioid carcinoma"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.1 kb. Library constructed by Life Technologies."
BASE COUNT 151 a 214 c 198 g 112 t
ORIGIN

Query Match 39.5%; Score 247; DB 141; Length 675;
Best Local Similarity 100.0%; Pred. No. 4.1e-40;
Matches 247; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 185 atgcgtgagaaactgacccgcgagagctcaagaagttcaagctgaagctctctcgg 244
|||||
DB 1 ATCGCTGAGAACTGACCGCCGAGAGCTCAAGAAGTTCAAGCTGCTGTCGG 60
QY 245 tgcgcgtgcgcgagggctacggcgccatcccgcgggcgctgctgtccatggagcgct 304
|||||
DB 61 TGGCGTGGCGAGGGCTACGGGCGCATCCCGGGGCGCTGCTGCCATGGACGCT 120
QY 305 tgaacctaccgacaagctgctgaagcttctacctgagacacacgagcgagctcacgc 364
|||||
DB 121 TGGACCTACCCACAAGCTGTGTCAGCTTCTACCTGGAGACCTACGGGCGCGAGCTCACCG 180
QY 365 ctacgtctgcgcgacatggcctgcagagagatggcgggcgagctgcagcgcgccacgc 424
|||||
DB 181 CTAACGTGTCGCGACATGGCGCTGCAGGAGATGCGGGCAGCTGCAGGCGCCACGC 240
QY 425 accaggg 431
|||||
DB 241 ACCAGGG 247

RESULT 9
AI346818/c 615 bp mRNA EST 02-FEB-1999
LOCUS qp58f05.x1 NCI_CGAP_C08 Homo sapiens cDNA clone IMAGE:1927233 3'
DEFINITION similar to TR:O15553 O15553 PYRIN. ; mRNA sequence.
ACCESSION AI346818
VERSION AI346818.1 GI:4084024
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 615)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
CONTACT: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.

cdNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio-llnl.gov/bbrp/image/image.html](http://bio-llnl.gov/bbrp/image/image.html)
Insert Length: 751 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 457.
Location/Qualifiers
1. 615
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:1927233"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from colon adenocarcinoma, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library is normalized. Library was constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 108 a 182 c 193 g 132 t
ORIGIN

Query Match 39.0%; Score 244.4; DB 19; Length 615;
Best Local Similarity 97.4%; Pred. No. 1.4e-39;
Matches 259; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 172 gacgccatctctgctgctgagaaacctgacccgcgagagctcaagaagttcaagctg 231
|||||
DB 615 GACGCCATCTTGGATCGCTGGAGAACCTGACCGCCGAGAGCTCAAGAAGTTCAAGCTG 556
QY 232 aagctgctgctgctgctgctgagggctacggcgccatcccgcgggcgctgctgctg 291
|||||
DB 555 AAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 496
QY 292 tccatgacgccttgacatccacgaagctggtgcagcttctacctgagacacacacgc 351
|||||
DB 495 TCCATGGACCCCTGGACCTCACCACAAAGCTGTGCTACCTGTGAGAGACCTACCGC 436
QY 352 gccagctccacgcctaacgtgctgcgcacatggggcctgcagagagatggcgggcgctg 411
|||||
DB 435 GCCGAGCTCACCGCTAACGCTGCTGCGGACATGGCGCTGCAGGAGATGCGCGGCACTG 376
QY 412 caggcgg-ccacgcacacagggtagc 436
|||||
DB 375 CAAGCGCCACCGCACCAGGGCTGC 350

RESULT 10
H16108 308 bp mRNA EST 27-JUN-1995
LOCUS Y128a06.r1 Soares breast 3NBHst Homo sapiens cDNA clone
DEFINITION IMAGE:159538 5', mRNA sequence.
ACCESSION H16108
VERSION H16108.1 GI:880928
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 308)
REFERENCE Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlff, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P. and Willson, R.
The WashU-Merck EST Project
JOURNAL Unpublished (1995)


```

mRNA sequence.
BE560228
VERSION BE560228.1 GI:9803948
KEYWORDS EST.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NIH-MGC http://mhc.nci.nih.gov/
1 (bases 1 to 830)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM356 row: d column: 08
High quality sequence stop: 728.
FEATURES
    source
    Location/Qualifiers
        1..830
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3678679"
            /clone.lib="NIH_MGC.8"
            /tissue_type="Burkitt lymphoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GCCACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 172 a 237 c 277 g 144 t

ORIGIN

```

Query Match 35.2%; Score 220.4; DB 137; Length 830;
Best Local Similarity 99.1%; Pred. No. 9e-35;
Matches 232; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 agcgctttgtgagggcaacgacggcgggggagtcgggagaccagagtgagg 60
|||||
Db 58 AGCGCCTTTTGTGAGGGCAACGGACCGGGCGGGAGTCGGGAGACCAAGTGGGAGG 117
|||||

QY 61 aagcggggagtcagggtccgcggcgagcgaacttcctctgtcggcggtgcagcg 120
|||||
Db 118 AAGGC-GGAGTCCAGGTTCGCCCGCGGAGCGACTTCCTCTGTCGGCGTGCAGCG 176
|||||

QY 121 ggtgagcgcgcagcggcgagtccttgagccatggggcgcgcgagcgcacatc 180
|||||
Db 177 GGGTAGCGCGCGCGAGCGCGGGGATCCCTGGAGCATGGGCGCGCGCGACGCCATC 236
|||||

QY 181 ctgtagctgtagaactgacccgcgagagagctcaagaagttcaagctgaagtg 234
|||||
Db 237 CTGGATGCTGCAGAACTGACCGCGGAGGAGCTCAAGAAGTTCAGACTGCAG 290
|||||

RESULT 13
BE560247
LOCUS BE560247 723 bp mRNA EST 15-AUG-2000
DEFINITION 60134605F1 NIH_MGC.8 Homo sapiens cDNA clone IMAGE:3678925 5',
mRNA sequence.
ACCESSION BE560247
VERSION BE560247.1 GI:9803967
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Ruminantia; Pecora; Bovidae;

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM356 row: n column: 14
High quality sequence stop: 649.
FEATURES
    source
    Location/Qualifiers
        1..723
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:3678925"
            /clone.lib="NIH_MGC.8"
            /tissue_type="Burkitt lymphoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GCCACGAG(G). Size-selected >500bp for average
            insert size 1.8kb. Library constructed by Ling Hong in
            the laboratory of Gerald M. Rubin (University of
            California, Berkeley) using ZAP-cDNA synthesis kit
            (Stratagene) and Superscript II RT (Life Technologies)."
```

BASE COUNT 136 a 219 c 256 g 112 t

ORIGIN

```

Query Match 34.9%; Score 218.6; DB 137; Length 723;
Best Local Similarity 84.8%; Pred. No. 2.1e-34;
Matches 243; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 1 agcgctttgtgagggcaacgacggcgggggagtcgggagaccagagtgagg 60
|||||
Db 58 AGCGCCTTTTGTGAGGGCAACGGACCGGGCGGGAGTCGGGAGACCAAGTGGGAGG 117
|||||

QY 61 aagcggggagtcagggtccgcggcgagcgaacttcctctgtcggcggtgcagcg 120
|||||
Db 118 AAGCGGGAGTCCAGGTTCGCCCGCGGAGCGACTTCCTCTGTCGGCGCTGCAGCG 177
|||||

QY 121 ggtgagcgcgcagcggcgagtccttgagccatggggcgcgcgagcgcacatc 180
|||||
Db 178 GGGTAGCGCGCGCGAGCGCGGGGATCCCTGGAGCATGGGCGCGCGCGACGCCATC 237
|||||

QY 181 ctgtagctgtagaactgacccgcgagagagctcaagaagttcaagctgaagctg 240
|||||
Db 238 CTGGATGCTGCAGAACTGACCGCGGAGGAGCTCAAGAAGTTCAGCTGCAGCGGCC 297
|||||

QY 241 tgggtgcgtgcgcaggggctacggggcgcatcccgcgggcgcgctgc 289
|||||
Db 298 AGCCACCAAGTGGCTCTTGGAGCCCCCGCGGCCCAAGGCTTGGGGCACTCC 346
|||||

RESULT 14
AV664060
LOCUS AV664060 555 bp mRNA EST 25-AUG-2000
DEFINITION AV664060 Bos taurus brain fetus Bos taurus cDNA clone E1BR040C01
5', mRNA sequence.
ACCESSION AV664060
VERSION AV664060.1 GI:9923090
KEYWORDS EST.
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

```

| | |
|----------------------|---|
| REFERENCE AUTHORS | Bovidae; Bovinae; Bos. 1 (bases 1 to 555) Sugimoto, Y., Hirotsune, S., Takasuga, A., Itoh, R., Jitchonzono, A. and Suzuki, H. |
| TITLE | bovine cDNA sequencing |
| JOURNAL | Unpublished (2000) |
| COMMENT | Contact: Yoshikazu Sugimoto Animal Genetics Division Shirakawa Institute of Animal Genetics Oadakra, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan Tel: 81-248-25-5641 Fax: 81-248-25-5725 Email: kazusugi@cocoa.ocn.ne.jp Single pass sequencing. This clone was obtained from a polyA-deleted cDNA library. |

```

FEATURES
    source
        Location/Qualifiers
            l..555
                /organism="Bos taurus"
                /db_xref="taxon:9913"
                /clone="E1BR040C01"
                /clone_lib="Bos taurus brain fetus"
                /tissue_type="brain"
                /dev_stage="fetus"
                /lab_host="DH10B"
                /note="Vector: pZLI; Site_1: SalI; Site_2: NotI; Poly A
                    was deleted from a NotI site"
            110 a 178 c 157 g 110 t
BASE COUNT
ORIGIN

```

| Query Match | 34.9%; | Score 218.4; | DB 32; | Length 555; |
|-----------------------|-----------------|--|-----------|-------------|
| Best Local Similarity | 83.0%; | Pred. NO. 2.3e-34; | | |
| Matches 249; | Conservative 0; | Mismatches 51; | Indels 0; | Gaps 0; |
| QY | 144 | ggatcctggagcctagggcgcgcgcgacgcacatcctgtgctgcgtcgagaaactgac | 203 | |
| Db | 15 | GGACCCCGGAGACATGGGGTGACACGCGGATGCCATCTTGGATGGCTGGGAGAACCTGAC | 74 | |
| QY | 204 | cgcgaggaagctcaagaagtcaagctgaagctgtctgttcgctgcgcgagaggcta | 263 | |
| Db | 75 | ACCGGACGAGCTCANAAGATTCAAGATGAAGTGTCTTTTCAGTGGCGTGTGGGAAAGGCTA | 134 | |
| QY | 264 | cgggcgcgcatcccgcgggcgcgctgtctgtccatggacgccttggacctcaccgacaagct | 323 | |
| Db | 135 | TGGACGCATCCCCCGGGGACCCTGCTGCCCTTGGATGCCGTGGACCTCACCGACAAGCT | 194 | |
| QY | 324 | gtaacgcttctaactgagacactacggcgcggaagctcaccgctaacgtgtgtgcgcgaat | 383 | |
| Db | 195 | CCTCAGCTACTCTCTGGAGGCGTACGGTGGCGAGCTACGGCGCTGTGTCTGGCGACAT | 254 | |
| QY | 384 | gggcctcgaggagatggcggggcagctgcaggcgccacgcaccagggtgagcgcgcccc | 443 | |
| Db | 255 | GGGCATGCAGGAGGTGGCGAGCAGCTGCAGAGAGACCATGACAAAGGTGAGCAGCCGCC | 314 | |

| | |
|------------|---|
| RESULT | 15 |
| AA442018 | |
| LOCUS | 446 bp mRNA EST 02-JUN-1997 |
| DEFINITION | zvw56c01.r1 Soares_total_fetus_ND2F8_9w Homo sapiens cDNA clone IMAGE:774048 5', mRNA sequence. |
| ACCESSION | AA442018 |
| VERSION | AA442018.1 |
| KEYWORDS | GI:2153896 |
| SOURCE | EST. |
| ORGANISM | human. |
| | Homo sapiens |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| REFERENCE | 1 (bases 1 to 446) |
| AUTHORS | Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Willson, R. |
| TITLE | WashU-Merck EST project 1997 |

JOURNAL
 COMMENT
 Unpublished (1997)
 Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information..
 Seq primer: -28m13 rev2 ER from Amersham
 High quality sequence stop: 385.
 Location/Qualifiers
 1..446
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:774048"
 /clone_lib="Soares_total_fetus_NB2HF8_9w"
 /dev_stage="8-9 weeks"
 /lab_host="DH108"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(dT) primer [5'
 TGTTCACCAATCTGAAGTCGGACGCGCTTAATTTTTTTTTTTTTTTT 3'].
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."
 79 a 143 c 162 g 59 t 3 others
 BASE COUNT
 ORIGIN
 Query Match 34.2%; Score 214; DB 7; Length 446;
 Best Local Similarity 85.6%; Pred. No. 1.7e-33;
 Matches 298; Conservative 0; Mismatches 37; Indels 13; Gaps 5;

| | | | | |
|-----------------------|-----------------|--------------------|------------|-------------|
| Query Match | 34.2% | Score 214; | DB 7; | Length 446; |
| Best Local Similarity | 85.6%; | Pred. No. 1.7e-33; | | |
| Matches 298; | Conservative 0; | Mismatches 37; | Indels 13; | Gaps 5; |

[illegible]

Search completed: August 30, 2001, 22:09:02
Job time: 5505 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2001, 20:42:32 ; Search time 39.27 Seconds
(without alignments)
3017.794 Million cell updates/sec

Title: US-09-691-763B-4
Perfect score: 626
Sequence: 1 agcgcctttgtgagggc.....ggaagggaagggaatcatt 626

Scoring table:
Gapex 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 52.6 | 8.4 | 44377 | 2 US-08-804-227C-7 | Sequence 7, Appli |
| 2 | 52.6 | 8.4 | 44377 | 2 US-08-804-198-1 | Sequence 1, Appli |
| 3 | 51.2 | 8.2 | 2943 | 1 US-08-042-747A-7 | Sequence 7, Appli |
| c 4 | 50.8 | 8.1 | 30001 | 1 US-08-125-468-1 | Sequence 1, Appli |
| c 5 | 50.8 | 8.1 | 30001 | 2 US-08-474-933-1 | Sequence 1, Appli |
| 6 | 50 | 8.0 | 2244 | 1 US-08-476-519-10 | Sequence 10, Appli |
| 7 | 50 | 8.0 | 2244 | 5 PCT-US95-09323-10 | Sequence 10, Appli |
| 8 | 50 | 8.0 | 2334 | 1 US-08-476-519-1 | Sequence 1, Appli |
| 9 | 50 | 8.0 | 2334 | 5 PCT-US95-09323-1 | Sequence 1, Appli |
| 10 | 49.4 | 7.9 | 1765 | 1 US-08-459-586-15 | Sequence 15, Appli |
| 11 | 49.4 | 7.9 | 1765 | 2 US-08-282-636-15 | Sequence 15, Appli |
| 12 | 49.4 | 7.9 | 18994 | 1 US-08-459-586-4 | Sequence 4, Appli |
| 13 | 49.4 | 7.9 | 18994 | 2 US-08-282-696-4 | Sequence 4, Appli |
| 14 | 48 | 7.7 | 530 | 3 US-08-758-662-4 | Sequence 4, Appli |
| 15 | 48 | 7.7 | 2310 | 3 US-09-036-987A-25 | Sequence 25, Appli |
| 16 | 48 | 7.7 | 4524 | 2 US-08-845-998-7 | Sequence 7, Appli |
| 17 | 48 | 7.7 | 4524 | 3 US-09-206-537-7 | Sequence 7, Appli |
| 18 | 47.4 | 7.6 | 800 | 5 PCT-US95-04801-4 | Sequence 4, Appli |
| c 19 | 47.4 | 7.6 | 50937 | 4 US-09-428-517-1 | Sequence 1, Appli |
| c 20 | 46.6 | 7.4 | 1656 | 4 US-09-385-028-14 | Sequence 14, Appli |
| 21 | 46.6 | 7.4 | 1809 | 1 US-08-455-001-1 | Sequence 1, Appli |
| 22 | 46.6 | 7.4 | 1809 | 5 PCT-US95-11869-1 | Sequence 1, Appli |
| c 23 | 46.6 | 7.4 | 15079 | 4 US-09-385-028-1 | Sequence 1, Appli |
| 24 | 46.4 | 7.4 | 4188 | 4 US-07-751-891B-2 | Sequence 2, Appli |
| 25 | 46.4 | 7.4 | 4242 | 3 US-07-705-490-2 | Sequence 2, Appli |
| 26 | 46.2 | 7.4 | 861 | 2 US-08-997-080-97 | Sequence 97, Appli |
| 27 | 46.2 | 7.4 | 861 | 2 US-08-997-362-97 | Sequence 97, Appli |

Sequence 97, Appli
Sequence 97, Appli
Sequence 7, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 23, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 1, Appli
Sequence 29, Appli
Sequence 2, Appli
Sequence 1, Appli

28 46.2 7.4 861 3 US-08-873-970-97
29 46.2 7.4 861 4 US-09-095-855-97
30 46 7.3 1506 4 US-09-158-767-7
31 46 7.3 2261 4 US-09-158-767-1
32 45.8 7.3 11219 1 US-07-642-734C-1
33 45.8 7.3 11219 3 US-08-439-009A-1
34 45.6 7.3 1227 4 US-09-385-028-23
35 45.6 7.3 15079 4 US-09-385-028-1
c 36 45.2 7.2 4692 2 US-08-916-917-1
c 37 45.2 7.2 4692 2 US-08-972-631-1
c 38 45.2 7.2 4692 2 US-08-972-629-1
c 39 45.2 7.2 4692 2 US-08-972-630-1
c 40 45.2 7.2 4692 2 US-08-972-629-1
c 41 45.2 7.2 4692 2 US-08-225-170-1
42 45 7.2 17612 3 US-08-911-853-29
43 45 7.2 17612 4 US-09-479-409-29
c 44 44.8 7.2 2580 3 US-09-050-863-2
45 44.8 7.2 5452 2 US-09-130-114-1

ALIGNMENTS

RESULT 1
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
; GENERAL INFORMATION:
; APPLICANT: DeHoff, Bradley S.
; APPLICANT: Rostock, Paul R., Jr.
; APPLICANT: Sutton, Kimberly L.
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: THOMAS G. PLANT 1501
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: ASCII(DOS) Text only
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/804.227C
FILING DATE: February 21, 1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Plant, Thomas, G.
REGISTRATION NUMBER: 35,784
REFERENCE/DOCKET NUMBER: X-8231
TELECOMMUNICATION INFORMATION:
TELEPHONE: 317-276-2459
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 44377 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 350..14002
FEATURE:
NAME/KEY: CDS
LOCATION: 14046..20036
FEATURE:
NAME/KEY: CDS
LOCATION: 20110..31284
FEATURE:

```
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-227C-7

Query Match      8.4%; Score 52.6; DB 2; Length 44377;
Best Local Similarity 55.9%; Pred. No. 0.023;
Matches 100; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 329 gcttctacctgagacacacgagcgccgagctacccgtacgctgctgcgacatgggccc 388
Db 16843 GCATCGACCTCCACCTACCTTCCAAACACGACCGCTACTGGCTGGCGCGCGTCCC 16902

QY 389 tgcaggagatggccgagcgtcagggccacccagggctgagccgcccccgctccc 448
Db 16903 GCCCCAGACGAGCGGCTGTGCGGCGCGGCTGTGCGGAGGTGGAGACCCCTGCTCA 16962

QY 449 cctccaccccgcttttccctccacccacacacagcgcttaccgccgggcttccgct 507
Db 16963 CCGCGCGGTGGAACCTGCCCGGACCGACGAGGTGTGGACCGCGGCATATCCGCT 17021

RESULT 2
US-08-804-198-1
; Sequence 1, Application US/08804198
; Patent No. 5945320
; GENERAL INFORMATION:
; APPLICANT: Burgett, Stanley G.
; APPLICANT: Kuhstoss, Stuart A.
; APPLICANT: Rao, Nagaraja R.
; APPLICANT: Richardson, Mark A.
; APPLICANT: Rosteck, Paul R., Jr.
; TITLE OF INVENTION: PLATENOLIDE SYNTHASE GENE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PAUL R. CANTRELL 1138
; STREET: LILLY CORPORATE CENTER
; CITY: INDIANAPOLIS
; STATE: IN
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/804,198
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CANTRELL, PAUL R.
; REGISTRATION NUMBER: 36,470
; REFERENCE/DOCKET NUMBER: P9113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3885
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44377 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 350..14002
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 14046..20036
; FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 20110..31284
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 31329..36071
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 36155..41830
; US-08-804-198-1

Query Match      8.4%; Score 52.6; DB 2; Length 44377;
Best Local Similarity 55.9%; Pred. No. 0.023;
Matches 100; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 329 gcttctacctgagacacacgagcgccgagctacccgtacgctgctgcgacatgggccc 388
Db 16843 GCATCGACCTCCACCTACCTTCCAAACACGACCGCTACTGGCTGGCGCGCGTCCC 16902

QY 389 tgcaggagatggccgagcgtcagggccacccagggctgagccgcccccgctccc 448
Db 16903 GCCCCAGACGAGCGGCTGTGCGGCGCGGCTGTGCGGAGGTGGAGACCCCTGCTCA 16962

QY 449 cctccaccccgcttttccctccacccacacacagcgcttaccgccgggcttccgct 507
Db 16963 CCGCGCGGTGGAACCTGCCCGGACCGACGAGGTGTGGACCGCGGCATATCCGCT 17021

RESULT 3
US-08-042-747A-7
; Sequence 7, Application US/08042747A
; Patent No. 5487969
; GENERAL INFORMATION:
; APPLICANT: Eberle, Richard
; APPLICANT: Scinicariello, Franco
; APPLICANT: Hilliard, Julia K.
; TITLE OF INVENTION: Cloning and Amplification of Monkey B
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cox & Smith Incorporated
; STREET: 112 East Pecan Street, Suite 2000
; CITY: San Antonio
; STATE: Texas
; COUNTRY: USA
; ZIP: 78205
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/042,747A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haymond, W. Bradley
; REGISTRATION NUMBER: 35186
; REFERENCE/DOCKET NUMBER: S-0072.179
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 210-554-5500
; TELEFAX: 210-226-8395
; TELEX: 767609
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2943 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: 87..2744
US-08-042-747A-7

Query Match 8.2% Score 51.2; DB 1; Length 2943;
Best Local Similarity 44.5%; Pred. No. 0.038;
Matches 203; Conservative 0; Mismatches 253; Indels 0; Gaps 0;

Qy 23 cggaccgggggagtcgggagaccagagtggaggaagcggggagtcacaggttcg 82
Db 208 CCGCCACGCCCGCTTCCCGAGGAGGAGGAGTCCCGAGGAGGAGGAGGAGGAGT 267

Qy 83 ccccgagcgcgactctctctggtcgcggtcgagcggtgagcggtcgagcggtcg 142
Db 268 CCACCCCGCGCGCCGCCAACGCGTGGTGGAGCGCGCGCGCGCGCGCGGAG 327

Qy 143 gggatcctggaacatggggcgcgcgagcgcacatcctggatgctgctggagaaactga 202
Db 328 ACCTGGGGAGATCAAGGCCCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGA 387

Qy 203 ccgcccggagagtcgaagaagtcgaagtcggtgctggtcggtcggtcggtcggt 262
Db 388 CCGCGCCACGGTGTGAGTTCGAGAGCGCCCGCGCGCGCGCGCGCGCGCGCG 447

Qy 263 acggcgcatcccgcggtcggtcggtcggtcggtcggtcggtcggtcggtcggt 322
Db 448 AGAATACACGAGGAGGATCGCGGTGCTTCAAGGAGAAATCGCCCGGTACAGTTCA 507

Qy 323 tggcagcttctacgtgagacctaagcggtcggtcggtcggtcggtcggtcggtcg 382
Db 508 AGGCCACCATGTACTACAGGAGCGTGCAGCGTCTGCTGCGCGCGCGCGCGCG 567

Qy 383 tgggctcgagagatggcggtcggtcggtcggtcggtcggtcggtcggtcggtcg 442
Db 568 CGCAGTTCATGGGATCTTCGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 627

Qy 443 cgttccctccaccgcgtcttctccctccaccacaca 478
Db 628 AGATCAACGCCAAGGGGTCTGCGCGGTGCGAGCGCCA 663

RESULT 4

US-08-125-468-1/c
; Sequence 1, Application US/08125468
; Patent No. 5589385
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracycline Formation and cosmid
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/125,468
; FILING DATE: 22-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsevdos, Estelle J
; REGISTRATION NUMBER: 31,145

REFERENCE/DOCKET NUMBER: 31,255-02
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)831-3241
TELEFAX: (201)831-3305
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 30001 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-125-468-1

Query Match 8.1% Score 50.8; DB 1; Length 30001;
Best Local Similarity 45.6%; Pred. No. 0.051;
Matches 219; Conservative 0; Mismatches 257; Indels 4; Gaps 1;

Qy 3 cgccttttgcggagggcaacgggagcgggggagtcgggagtcgggagtcgggagtc 62
Db 24497 CGCGGGTTCAGAGCGGATCGTCTCGCGCGAGCGCGGAGTGGAGCGGGATGACCGAC 24438

Qy 63 ggcggaggtccaggttcgccccggagcgacttctctctggtcggtcggtcggtcg 122
Db 24437 GTCTGTGTGACCGGATCGGC----GTGGTGGCACCGACGGCTGGGGCTTGGGAGTA 24382

Qy 123 gtgagcggtcggtcggtcggtcggtcggtcggtcggtcggtcggtcggtcggt 182
Db 24381 CTGGGAGCGCACATGCGCGCGCGGTCTCGGCTCTCGCGCGCGGTCTCGACTCCAC 24322

Qy 183 ggtgctgctggaacactgaacggcgagagtcgaagaagttcaagtcggtgctg 242
Db 24321 CGGTACCGGGCGAGTGGCGCGGAGGTGAATTCACGCCGCGGAGAACACCTGCCGG 24262

Qy 243 ggtgctgctggtcggtcggtcggtcggtcggtcggtcggtcggtcggtcggt 302
Db 24261 CCGACTGCTCGCGCGAGACCGACATGACGCGGCTGCGCGCGCGCGCGGAGG 24202

Qy 303 ctggtgactcaccgacaaagtgtcagcttctctctggtgagacctacggtcggtcg 362
Db 24201 GCTCGCGCGCGGGGCGGACCGCGACGAGTCTCGCGGCTTCTCGCGCGCGGTGAC 24142

Qy 363 cgttaacgtgctgctgacatggtcggtcggtcggtcggtcggtcggtcggtcggt 422
Db 24141 GCGCGCTCGCGCGCGGCTTGGAGTTCGCGCGAGAGAACTCCAGGCGCTGTGGAGCAA 24082

Qy 423 gcacaggtgagcggtcggtcggtcggtcggtcggtcggtcggtcggtcggtcg 482
Db 24081 GGGCGCGCGCTACGTGACGCGCTACAGTCTCGGTGCTTCTACCGGTCAACACGG 24022

RESULT 5

US-08-474-933-1/c
; Sequence 1, Application US/08474933
; Patent No. 5866410
; GENERAL INFORMATION:
; APPLICANT: Ryan, Michael J.
; APPLICANT: Lotvin, Jason A.
; APPLICANT: Strathy, Nancy
; APPLICANT: Fantini, Susan E.
; TITLE OF INVENTION: Cloning of the biosynthetic pathway for
; TITLE OF INVENTION: chlortetracycline and tetracycline Formation and cosmid
; TITLE OF INVENTION: useful therein
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESS: American Cyanamid Company
; STREET: One Cyanamid Plaza
; CITY: Wayne
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07470
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

| | | | |
|----|-----|---|-----|
| Db | 104 | CGTAGGGCGCCGCGGGGGCGCGATTTGGCTGCTTAACGGCGAGCGGCACAGCGGTGG | 163 |
| QY | 83 | ccccgagcagcagcttctctctggttcggcggcctgcagcggggtgagcggcgagcagcgccg | 142 |
| Db | 164 | TGCGGGCGAGCGCGCGACACAGCAGCGCGCGCCACGGTGATGGTGGGTTTCGACAGCGATGG | 223 |
| QY | 143 | gggatcctgagagcattggggcgcgcgcgacgcacatccttgatgcgctgagagaaacctga | 202 |
| Db | 224 | CGGGTCCGGCGCGGTGCGGGACCGTGC CGGCGGCCATGCTGGTGGTGCTGCTGGACGAGC | 283 |
| QY | 203 | ccgcgcgagagctcaagaagttcaagctgaagctgctgcggtgcgcctgcgcgagggct | 262 |
| Db | 284 | TGGGCGCGGTGTTTCGGGTACTGCCCGCTGAGCGGCGACGGTGATCCCGCTGGCGCGGAGC | 343 |
| QY | 263 | acgggcgcattccgcgggggcgcgtgctgtccattgagcgccttggacctcacccacaagc | 322 |
| Db | 344 | TGTCGCATCTTCTGCGCGCGGGCGTGCTGGCGCGCTGGCGCTGGGACCGCAGTCTGGCGC | 403 |
| QY | 323 | tggtcaagctcttaccttgagacctcaaggcgcgagagctcaacgctaaagtgctgcgcgaca | 382 |
| Db | 404 | CCGCGCGCGAGCGCGCGGGGGTGTGCTCCCGAGCTGGACCGCGAGTGGGAGCGCGC | 463 |
| QY | 383 | tgggcctgcagagagatggtccggggcagctgcagggcgccagccaccaggggtgagcgccc | 441 |
| Db | 464 | CGCGTGGGACGGGTTCACCTGCACCGCGCGCGCGCTGTGTGGGCGCGCAGCGCAGC | 522 |

```

RESULT 11
US-08-283-696-15
: Sequence 15, Application US/08282696
: Patent No 5845806
: GENERAL INFORMATION:
: APPLICANT: Jones, Thomas R.
: APPLICANT: Campbell, Ann E.
: TITLE OF INVENTION: Identification of a Human
: TITLE OF INVENTION: Cytomegalovirus Gene Region Involved
: TITLE OF INVENTION: MHC Class I Heavy Chain Expression
: TITLE OF INVENTION:

```

Query Match 7.9%; Score 49.4; DB 2; Length 1765;
Best Local Similarity 44.9%; Pred. No. 0.083;

Matches 188; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

Qy 23 cggaccggggcgggagdtcgggagaccagtgaggagaagcgggagtgccaggtccy 82
Db 104 CGTACGGGCGCGCGGGGGGGCGGATTTGGCTGCCTTACGCGGAGCGCGGCGGGTGG 163
Qy 83 ccccgagccgacttctctctggtcgggcggtcgagcggggtgagcgggcgagcgggcg 142
Db 164 TCGGGGCGGACGCGGACAGCAGCGCGCGGCGGCGGTCGATGTTGGTGGTTCGACAGCATGG 223
Qy 143 gggatccctggagccatggggcgcgcgagcgcacccatctgagtcgctggagaacctga 202
Db 224 CGGGTCGGGCGGCTCGGGGAGACGTGTCGGCGGCCCTGCTGGTGGTCTGCTGGAGAGAC 283
Qy 203 cgcgcgagagctcaagaagtccaagctgaagctgctgtcggtgcgcgtgcgcgaggtc 262
Db 284 TGGGCGCGCGTGTTCGGGTACTTGGCCGCTCGAGCGGCGACGTTGTACCCGCTGGCGGGGAGC 343
Qy 263 acggcgcatcccgcgggggcgcgctgctgctccatgagcgcgcttgagcctcaccgacaagc 322
Db 344 TGTGCACATTTCTGCGCGCGGGGTGCTGGGGCGCGTGGCGCTGGGACGCGAGTCGGCGC 403
Qy 323 tggtaagcttctacctagacacctaagcgcgcgagctcaccgtacgtgctgcgcgaca 382
Db 404 CGCGCGGAGCGCGCGGGGCTGCTCCCGAGACTGGACCGCGGACGAGTCGGAGCGCGC 463
Qy 383 tgggctcgagagatgcccggcgagctcgagcgcgccagcgaacaggtgagccgcc 441
Db 464 CGCGTGGGACGCGCTGCACCTGCACCGCGCGCGCGCTGTGGGCGCGGAGCGCGAC 522

RESULT 12
US-08-459-586-4
; Sequence 4, Application US/08459586
; Patent No. 5720957
; GENERAL INFORMATION:
; APPLICANT: Jones, Thomas R.
; APPLICANT: Campbell, Ann E.
; TITLE OF INVENTION: Identification of a Human
; TITLE OF INVENTION: Cytomegalovirus Gene Region Involved in Down-Regulation of
; TITLE OF INVENTION: MHC Class I Heavy Chain Expression
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESS: American Cyanamid Company
; STREET: One Campus Drive
; CITY: Parsippany
; STATE: New Jersey
; COUNTRY: United States
; ZIP: 07054
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/459,586
; FILING DATE: 02-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Barnhard, Elizabeth M.
; REGISTRATION NUMBER: 31,088
; REFERENCE/DOCKET NUMBER: 32,425-01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-683-2158
; TELEFAX: 201-683-4117
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 18994 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-459-586-4

LENGTH: 530 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-758-662-4

Query Match 7.7%; Score 48; DB 3; Length 530;
 Best Local Similarity 55.4%; Pred. No. 0.15;
 Matches 93; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 23 cgacccggcgagctcggagaccagagtggaagagcggggagtcaggttcg 82
 Db 310 CGCGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 369
 Qy 83 ccccgagccgactctctctgtgcggtgcagcggggtgagcgggcgagcgccg 142
 Db 370 CCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG 429
 Qy 143 gggatctgagccatggcgcgcgcgagccatctgtgatgcg 190
 Db 430 GGGGGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCG 477

RESULT 15

US-09-036-987A-25
 ; Sequence 25, Application US/09036987A
 ; Patent No. 6143526

GENERAL INFORMATION:

APPLICANT: Baltz, Richard H.
 APPLICANT: Broughton, Mary C.
 APPLICANT: Crawford, Kathryn P.
 APPLICANT: Madduri, Krishnamurthy
 APPLICANT: Merlo, Donald J.
 APPLICANT: Treadway, Patti J.
 APPLICANT: Turner, Jan R.
 APPLICANT: Waldron, Clive
 TITLE OF INVENTION: Biosynthetic Genes For Spinosyn Insecticide
 TITLE OF INVENTION: Production
 NUMBER OF SEQUENCES: 39
 CORRESPONDENCE ADDRESS:

ADDRESSEE: Dow Agrosciences LLC Patent Department
 STREET: 9330 Zionsville Road
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: USA
 ZIP: 46268

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/036,987A
 FILING DATE: 09-MAR-1998
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stuart, Donald R.
 REGISTRATION NUMBER: 28,479
 REFERENCE/DOCKET NUMBER: 50,608
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (317)337-4816
 TELEFAX: (317)337-4847

INFORMATION FOR SEQ ID NO: 25:

SEQUENCE CHARACTERISTICS:
 LENGTH: 2310 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 88..1077

FEATURE:
 NAME/KEY: CDS
 LOCATION: 1165..1992
 US-09-036-987A-25

Query Match 7.7%; Score 48; DB 3; Length 2310;
 Best Local Similarity 48.5%; Pred. No. 0.15;
 Matches 132; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

Qy 81 cgccccggagccgactctctctgtgcggtgcagcggggtgagcgcgagcgccg 140
 Db 516 CGCGCTGGAGCCGAATTCCTCCGTACTCGCGGGCGAAGCGGCTCGGACCTGCTGGCCG 575
 Qy 141 cggggatctgagccatggggcgcgcgagccatctctgtgatcgctgagaaacct 200
 Db 576 CGCCTACCAACCGCACCCACGGACTGCCGGTGTGCATCACCCGCTCTCCAACAACACG 635
 Qy 201 gacgcgcgagggagctcaagaagttcaagctgaagctgctcggtgcgctgcgcgaggg 260
 Db 636 GCCCTACCAAGTTCCTCCGGAGAGGTGCTGCCGCTGTTTCATCAGAACCTGATGGACGGCAG 695
 Qy 261 ctacggcgcatcccgcgggcgcgctgtgtccatgagcgcttgacccctcaccacaa 320
 Db 696 CCAGGTGCCGCTCTACGGCGACGGGCTCAACGTGGCGGACTGGCTGCACGCTCAGCGACCA 755
 Qy 321 gctggctcagcttctacctggagacctacggcg 352
 Db 756 CTGCGGGGCGCATCCAGTGTGGCGGACTCCG 787

Search completed: August 30, 2001, 22:31:18
 Job time: 6526 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2001, 21:14:47 ; Search time 81.93 seconds
(without alignments)
4797.588 Million cell updates/sec

Title: US-09-691-763B-4
Perfect score: 626
Sequence: 1 agcgctttgtgagggc.....ggagggagggatcactt 626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues
Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-----------------------------|
| 1 | 389 | 62.1 | 811 | 21 | AAC77884 Human cancer assoc |
| 2 | 368 | 58.8 | 779 | 20 | AAZ33631 Human breast tumou |
| 3 | 322 | 51.4 | 740 | 22 | AAF30007 Human CARD-5 cDNA |
| 4 | 320 | 51.1 | 432 | 21 | AAZ98638 Human colon cancer |
| 5 | 178.8 | 28.6 | 777 | 22 | AAF30008 Mouse CARD-5 cDNA |
| 6 | 174.4 | 27.9 | 807 | 22 | AAC88099 Human FLEXHT-30 nu |
| 7 | 86.2 | 13.8 | 16891 | 20 | AAZ37084 MFV gene sequenced |
| 8 | 74.8 | 11.9 | 3512 | 20 | AAZ37085 FMF associated pro |
| 9 | 68.4 | 10.9 | 6461 | 22 | AAZ00033 Human ATLAS-2-enco |
| 10 | 52.6 | 8.4 | 44377 | 18 | AAZ78508 Platenolide syntha |
| 11 | 52.6 | 8.4 | 44377 | 18 | AAZ80414 Platenolide syntha |

| | | | | | | | |
|---|----|------|-----|--------|----|----------|---------------------|
| c | 12 | 51.8 | 8.3 | 37856 | 21 | AAAL1992 | S. cellulosum DNA |
| c | 13 | 50.8 | 8.2 | 2943 | 17 | AAT16480 | SA8 virus gB glyco |
| c | 14 | 50.8 | 8.1 | 30001 | 18 | AAT61016 | Total DNA sequence |
| c | 15 | 50.8 | 8.1 | 30001 | 21 | AAZ05110 | S. aureofaciens DN |
| c | 16 | 50.6 | 8.1 | 11279 | 21 | AAZ38389 | Pseudomonas sp. WF |
| c | 17 | 50.6 | 8.1 | 53789 | 19 | AAZ21187 | Anycolatopsis medi |
| c | 18 | 50 | 8.0 | 2244 | 17 | AAT10429 | Mature isoamylase |
| c | 19 | 50 | 8.0 | 2244 | 19 | AAZ33640 | Flavobacterium iso |
| c | 20 | 50 | 8.0 | 2334 | 17 | AAT10428 | Isoamylase gene. |
| c | 21 | 50 | 8.0 | 2334 | 19 | AAZ33639 | Full length Flavob |
| c | 22 | 49.4 | 7.9 | 1765 | 20 | AAV08427 | HCWV strain AD169 |
| c | 23 | 49.4 | 7.9 | 18994 | 20 | AAV08416 | HCWV strain AD169 |
| c | 24 | 48.8 | 7.8 | 1925 | 20 | AAZ90924 | Epstein Barr Virus |
| c | 25 | 48.4 | 7.7 | 2061 | 10 | AAZ92408 | Sequence encoding |
| c | 26 | 48 | 7.7 | 530 | 21 | AAZ5215 | Human OCT-TL parti |
| c | 27 | 48 | 7.7 | 1272 | 20 | AAZ09010 | Brn-3a polynucleot |
| c | 28 | 48 | 7.7 | 1272 | 21 | AAZ09006 | Human transcriptio |
| c | 29 | 48 | 7.7 | 2310 | 20 | AAZ21502 | Saccharopolyspora |
| c | 30 | 48 | 7.7 | 4524 | 20 | AAZ33912 | Nucleotide sequenc |
| c | 31 | 48 | 7.7 | 114955 | 20 | AAZ53491 | Human adenosine A1 |
| c | 32 | 47.8 | 7.6 | 2538 | 20 | AAZ21398 | Central fragment o |
| c | 33 | 47.4 | 7.6 | 2672 | 16 | AAT08693 | Cephameycin biosynt |
| c | 34 | 47.4 | 7.6 | 50837 | 21 | AAZ09469 | Streptococcus olea |
| c | 35 | 47 | 7.5 | 1844 | 12 | AAZ12000 | Maize nitrite redu |
| c | 36 | 46.6 | 7.4 | 1809 | 17 | AAT15008 | Mouse Elf-1 cDNA. |
| c | 37 | 46.6 | 7.4 | 1809 | 19 | AAZ42926 | CDNA encoding a ma |
| c | 38 | 46.6 | 7.4 | 15079 | 16 | AAQ91580 | S. Clavuligerus cl |
| c | 39 | 46.4 | 7.4 | 2106 | 20 | AAZ28103 | Freac11 gene. Hom |
| c | 40 | 46.2 | 7.4 | 861 | 19 | AAZ34602 | M. vaccae antigen |
| c | 41 | 46.2 | 7.4 | 861 | 20 | AAZ11337 | M. vaccae potb gen |
| c | 42 | 46 | 7.3 | 1506 | 20 | AAZ36871 | Codon modified cDN |
| c | 43 | 46 | 7.3 | 2261 | 20 | AAZ36865 | Cytochrome P450 cD |
| c | 44 | 45.8 | 7.3 | 1803 | 21 | AAZ22369 | Human secreted pro |
| c | 45 | 45.8 | 7.3 | 1845 | 21 | AAZ98400 | Canine beta1-adren |

ALIGNMENTS

RESULT 1
AAC77884
ID AAC77884 standard; CDNA; 811 BP.
AC AAC77884;
DT 08-FEB-2001 (first entry)
XX Human cancer associated gene sequence SEQ ID NO:278.
DE Human; cancer associated gene; cancer antigen; detection; cancer;
KW diagnosis; cytostatic; proliferative; vulnery; immunomodulator;
KW antidiabetic; antisthmatic; antirheumatic; antiallergic; antiviral;
KW antiinflammatory; antihypertensive; antiallergic; antibacterial; cardiant;
KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
KW allergic reaction; graft versus host disease; organ rejection;
KW haemostatic; thrombolytic; cardiovascular disorder; infection;
KW neurological disease; drug screening; ss.
XX Homo sapiens.
OS WO2000055350-A1.
XX PN 21-SEP-2000.
XX PD 08-MAR-2000; 2000WO-US05882.
XX PF 12-MAR-1999; 99US-012470.
XX PR (HUMA-) HUMAN GENOME SCI INC.
XX PA Rosen CA, Ruben SM;
XX PI


```

XX Dehoff BS, Kuhstoss SA, Rostock PR, Sutton KL;
PI WPI; 1997-418046/39.
XX P-PSDB; AAW22606-W22610.
DR DNA encoding Streptomyces fradiae tyactone synthase domain - for
XX production of tylosin-related polyketide compounds
XX Example 2; Pages 110-134; 220pp; English.
XX
CC This sequence represents the platenolide synthase gene cluster of the
CC invention. This sequence is referred to as the smg gene, and was
CC isolated from Streptomyces ambifaciens. This sequence encodes the
CC multi-functional proteins which direct the synthesis of the polyketide
CC platenolide. Platenolide is the basic building block of the macrolide
CC antibiotic spiramycin. This sequence was used along with the tyIG gene
CC (see AAT80413) to create a hybrid ORF1 sequence (see AAT80415). The tyIG
CC gene is the tyactone synthase gene cluster of the invention. The tyIG
CC sequence was isolated from Streptomyces fradiae, and encodes
CC multifunctional proteins which direct the synthesis of the polyketide
CC tyactone. Tyactone is the basic building block of the antibiotic
CC tylosin. The hybrid sequence can be used to transform S. ambifaciens
CC lacking the smg ORF1 sequence, or S. fradiae lacking the tyIG ORF1
CC sequence, so that they can produce polyketides. The DNA sequence can be
CC modified so as to alter the type of carboxylic acids incorporated, the
CC number of carboxylic acids incorporated and/or the post-condensation
CC reactions performed, thereby resulting in novel tylosin-related
CC polyketides.
XX
XX Sequence 44377 BP; 4965 A; 15552 C; 17381 G; 5479 T; 0 other;
SQ
Query Match 8.4%; Score 52.6; DB 18; Length 44377;
Best Local Similarity 55.9%; Pred. No. 0.15;
Matches 100; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
Oy 329 gctctacgtgagactgagcgccgagctcaccgtacgtctgctgcgacatgggcc 388
Db 16843 gcatcgactccctaccctacccttccacacgacgctactgctggcgccctcc 16902
Oy 389 tgcagagatggcgggcgagctgacggcgccacgacgaggtgagccgccccttcc 448
Db 16903 ggcacgagcgagcggtgtgctgctgctgctgctgctgctgctgctgctgct 16962
Oy 449 cctccacccgcttctccctccctccacacacgacgcttaccgctgcttccgct 507
Db 16963 ccgcccgtggaactgcccgccacgacacgaggtgtgacccgcccgcataccgct 17021
RESULT 12
AA11992/c
ID AA11992 standard; DNA; 37856 BP.
XX
XX AA11992;
XX
XX 07-AUG-2000 (first entry)
XX
XX S. cellulosum DNA encoding polyketide and hereropolyketide enzymes.
XX
XX Polyketide biosynthesis; heteropolyketide biosynthesis; mutasynthesis;
XX epothilone; cytotoxic; immunosuppressant; antibiotic; antifungal;
XX plant-protection; ds.
XX
XX Sorangium cellulosum.
XX
XX Key Location/Qualifiers
XX CDS Complement (3398..6100)
XX FT /*tag= a
XX FT /*product= "ORF1-trna synthetase"
XX FT /*note= "gtg start codon"
XX FT complement (6374..7111)
XX FT /*tag= b

```

```

FT CDS
FT complement (8433..9550)
FT /*tag= c
FT /*product= "ORF3-amino transferase"
FT /*note= "AGT start codon given in the specification"
FT 9855..11393
FT /*tag= d
FT /*product= "ORF4- tyrosine/DOPA-Decarboxylase"
FT /*note= "GTG start codon"
FT 12212..13658
FT /*tag= e
FT /*product= "ORF5-3-oxoacyl-ACP-reductase"
FT /*note= "ACC start codon"
FT 15374..119984
FT /*tag= f
FT /*product= "ORF6-polyketide synthase"
FT 20003..27889
FT /*tag= g
FT /*product= "ORF7-peptide synthetase"
FT 28251..29400
FT /*tag= h
FT /*product= "ORF8-transpeptidase"
FT complement (30040..31720)
FT /*tag= i
FT /*product= "ORF9-regulation element"
FT /*note= "CGC stop codon"
FT 31982..32932
FT /*tag= j
FT /*product= "ORF10-transcription regulator"
FT 33128..33613
FT /*tag= k
FT /*product= "ORF11-regulation element"
FT /*note= "GTG start codon"
FT 33661..34077
FT /*tag= l
FT /*product= "ORF12-regulation element"
FT complement (35255..35616)
FT /*tag= m
FT /*product= "ORF13-transcription regulator"
FT complement (35730..36242)
FT /*tag= n
FT /*product= "ORF14-transcription regulator"
FT /*note= "GTG start codon"
DE19846493-Al.
XX
XX 13-APR-2000.
XX
XX 09-OCT-1998; 98DE-1046493.
XX
XX 09-OCT-1998; 98DE-1046493.
XX
XX (GBFB ) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
XX
XX Beyer S, Mueller R;
XX
XX WPI; 2000-294101/26.
XX
XX DNA sequence coding for products involved in the biosynthesis of
XX polyketide or heteropolyketide compounds, especially epothilone
XX Claim 3; Page 20-33; 36pp; German.
XX
XX This invention describes a novel DNA sequence (I) whose expression
XX products effect or are involved in the enzymatic biosynthesis,
XX mutasynthesis or partial synthesis of polyketide or heteropolyketide
XX compounds (II). (I) can be inserted into an expression vector and used
XX to transform or transfect prokaryotic or eukaryotic cells with the aim
XX of obtaining strains that produce large amounts of polyketide or
XX heteropolyketide compounds, especially epothilones, which have cytotoxic
XX and/or immunosuppressant and antibiotic and antifungal activities and
XX are useful as plant-protection agents. This sequence represents the DNA
XX sequence isolated from Sorangium cellulosum which is described in the

```

CC method of the invention.

```
XX
XX
SQ Sequence 37856 BP; 5655 A; 13666 C; 12913 G; 5622 T; 0 other;

Query Match      8.3%; Score 51.8; DB 21; Length 37856;
Best Local Similarity 47.7%; Pred. No. 0.21;
Matches 218; Conservative 0; Mismatches 232; Indels 7; Gaps 2;

QY 4 gcttttgcagaggaacgacgggaggggagtgaggagaccagagtgaggaggaag 63
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31674 GCCGCTGGCGAGGAGCAAGCCCTCAAGCTCTCGAGGACATCGCTGCCCGCGAGG 31615

QY 64 gcggggagtcagagttccgcccggagcgagcttctctgtgctggcggtcagcggg 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31614 TGGAGAGCGCTTCTTCGCGCGCGGAGGACCGCTCCCGAGGTGAGTACCGGTGCG 31555

QY 124 tgagcgcgagcgagcgagcgagctgagccatgggagcgagcgagcgagcgagc 183
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31554 ATCCGAGCGGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 31499

QY 184 gatcgctggagaacctgacccgagggagctcaaggttcaagctgaagctgctgtcg 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31498 GACGGGAGCGCGCGGCGGCTCGGCTGGCTGGCGGACACGTCGCGCGGAGATCCAGGCC 31439

QY 244 gtgcgctgctgcgagggactacggcgcatcccgggcgcgct--gctgtccatgac 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31438 GCGGAGCTGCTCGAGGCGCGGCGACCCCGCGCTTCGCGCGGCTCGGAGAGCTCTAC 31379

QY 301 gcttggacctcaacgagaagctggtcagcttctacctggagacctacggcgagctc 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31378 GCGCGCGCGCGCTCGCGGTTCTTCGCGCGGCTCGCTCAAGAACATCGACCTCGCGGAGCAC 31319

QY 361 accgctaactgtgctgcagacatgggctcgcagagatggcgggcagctgcagcgacc 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31318 CTCACGAGCGCTGCGCGTCCACGCTGGGACGAGCGGACGATCCGAGGAGGAGCGG 31259

QY 421 acgacacagggtagcgcccccggttccctccacc 457
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31258 CTCGACGCGCGCGCTTCGCGGACATGCTCGCGCGGCC 31222

RESULT 13
AAT16480
ID AAT16480 standard; DNA; 2943 BP.
XX
AC AAT16480;
XX
DT 11-MAY-1996 (first entry)

SA8 virus gB glycoprotein coding sequence.
XX
DE Herpes simian monkey SA8 virus gB glycoprotein; immunoassay;
KW diagnosis; herpes B virus; ss.
XX
OS Herpes simian monkey SA8 virus.
XX
FH Key Location/Qualifiers
CDS 87..2744
FT /*tag= a
FT /product= gb_glycoprotein
XX
XX US5487969-A.
XX
XX 30-JAN-1996.
XX
XX 01-APR-1993; 93US-0042747.
XX
XX 01-APR-1993; 93US-0042747.
XX
XX (SWBI-) SOUTHWEST FOUND BIOMEDICAL RES.
XX
XX Black D, Eberle R, Hilliard J, Scinicariello F;
```

```
XX
DR WPI; 1996-105220/11.
XX P-PSDB; AAR92747.
XX
PT Detection of herpes B virus by PCR amplification of sample DNA - to
XX detect a specific herpes simian monkey B virus DNA segment.
XX
PS Disclosure; Column 23-30; 22pp; English.
XX
CC This sequence encodes the herpes simian monkey SA8 virus gB
CC glycoprotein. Such sequences have been used to develop synthetic
CC DNA primers which have homologous sequences of conserved regions
CC which flank a divergent region of the gB glycoprotein gene. In
CC addition, proteins such as the gB glycoprotein have potential use in
CC the development of serological immunoassays. One approach is to
CC synthesize peptides which, based on the properties of the predicted
CC protein sequence, are likely to be immunologically active. Such
CC peptides can be used as substrate antigens in immunoassays to detect
CC serum antibodies which recognize this specific peptide sequence.
CC Synthetic peptides may also be used to produce antibodies against
CC specific regions of the gB glycoprotein which are unique to one
CC virus. These can then be used to develop virus-specific
CC immunoassays for differentiation of SA8 virus from other primate
CC alpha-herpes viruses and for identification of antibodies directed
CC against SA8 virus in primate serum samples.
XX
SQ Sequence 2943 BP; 462 A; 1103 C; 983 G; 395 T; 0 other;

Query Match      8.2%; Score 51.2; DB 17; Length 2943;
Best Local Similarity 44.5%; Pred. No. 0.28;
Matches 203; Conservative 0; Mismatches 253; Indels 0; Gaps 0;

QY 23 cggaccggggggagtcgggagacacagtgaggagagcgaggcggtccaggttccg 82
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 cgcacacggcgcttcccgaggagagaggtcccgagagacggaggggggtcg 267

QY 83 ccccgagcgagcttctctgtgctggcggtcagcggtgagcgggcgagcgagcgcg 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 268 ccacccggcgcccgcccaacgctcggtcgagggcgcgcgagcgctgcggaag 327

QY 143 gggatcctggagccatggcgcgcgcgagccatcctggtgctgctgagaaactga 202
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 328 acctcgggagatacaaggccgggagcgagcgaccttctacgtctgcccgcgcga 387

QY 203 ccgcgagagagctcaagaagtcaagctgaagctgctgctggtgctgctgagggct 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 388 cggcgccacggtggtgagtcgagcagcccgcggtgcccgcgcgcgcgcgcgc 447

QY 263 acgggcgcatcccgggcgcgctgctgctcatgagcgcttgacctcaccgacaagc 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 448 agaaactacacggaggggtcggtgcttcaagagagaacatgcccgtacaagtca 507

QY 323 tggtaagttctacctggagacctacggcgcgagctacccgttaacgtgctgcgcga 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 aggccaccatgtactacaagagctgacctgtctcaggtgtgttctcggcagcggt 567

QY 383 tgggctcgagagatggcgggcgagctgcagcgcccgagccagccaggtgagccccc 442
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 568 cgcagttcatggggtctctcagggagcccgcccccgctcccttcgagaggtgatgaca 627

QY 443 cgttccccctcacccccgttcttccccctcacccaca 478
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 628 agatcaacgccaagggggtgtgctgctgagcgagcgcca 663

RESULT 14
AAT61016/C
ID AAT61016 standard; DNA; 30001 BP.
XX
AC AAT61016;
XX
DT 21-APR-1997 (first entry)
```


Qy 3 cgcccttttgcgagggcaacgagccggggcggggagtcgaggagaccagagtgaggagaa 62
 Db 24497 CGGCGGCTTCCAGAGCGCGATCGTGTCCGAGCGCGAGTGGAGCGGGGATGACCGAC 24438
 Qy 63 ggcgggagtcgaagttcccccgcgcgagccgaacttcctcctggtcggcgctgcagcggg 122
 Db 24437 GTCTGGTACCGGGAFCGGC----GTGTGGCACCGAACCGCTGGGGCTTCGGAGTA 24382
 Qy 123 gtgagcgcgagcgccggggtatccttgagccatggggcgcgcgagcgccatcct 182
 Db 24381 CTGGGACGCCACACTGGCGGCGCGGTCCGGTCTCGGGCGCTGACCGCGCTTCGACTCCAC 24322
 Qy 183 gpatgcgctgagaaacctgaccgccgagagctcaagaagttcaagctgaagctctctc 242
 Db 24321 CGGGTACCGGGCGCACGTGGCGCGGAGGTGAACCTTCACGCCGGAGGAACCTGCCGGG 24262
 Qy 243 ggtgcgctgcgaggggctacggcgccatcccgcgggcgcgctgctgtccatggaagc 302
 Db 24261 CCGACTGTCTGCCGAGACCGACACATGACGGGGCTGGCCCTGTCTCGGGCGCGAGAGGC 24202
 Qy 303 cttggacctcaccgacaaagctggtcagcttctacctggagacctacggcgcgagctcac 362
 Db 24201 GCTCGCGACGCGGGGCGGACCGGAGACGATGCCGGACTTCTCCCGCGGTCTGTGAC 24142
 Qy 363 cgctaactgtgcgcgacatgggacctgcagagatggccggcgagctgcagcgggccac 422
 Db 24141 GGCCTGCTCGCGCGCGGCTTCGAGTTCGGCCAGAAAGAACTCCAGCGGCTGTGGAGCAA 24082
 Qy 423 gcaccagggtgagccgcccccggtcccccctccaccgcgtcttccctccaccaccacag 482
 Db 24081 GGGCCGGGCTACGTACGCGCTTACCACTCTTCCGCTGTCTTACCCGGTCAACACGG 24022

Search completed: August 30, 2001, 22:32:50
 Job time: 4683 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 30, 2001, 20:41:57 ; Search time 1242.53 Seconds
(without alignments)
7792.819 Million cell updates/sec

Title: US-09-691-763B-4
Perfect score: 626
Sequence: 1 agcgcttttctgagggc.....ggaagggaaggatcactt 626

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_viz:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pl:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rod:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|---------------|-------|----------------|--------|----|--------------------|
| 1 | 626 | 100.0 | 2821 | 89 | AF184072 Homo sapi |
| 2 | 626 | 100.0 | 233425 | 60 | AC009088 Homo sapi |
| 3 | 368 | 58.8 | 779 | 9 | AX017270 Sequence |
| 4 | 361 | 57.7 | 782 | 85 | AB023416 Homo sapi |
| 5 | 342.8 | 54.8 | 713 | 89 | AF255794 Homo sapi |
| 6 | 342.6 | 54.7 | 770 | 89 | AF184073 Homo sapi |
| 7 | 322 | 51.4 | 740 | 10 | AX082246 Sequence |
| 8 | 301 | 48.1 | 740 | 89 | AK000211 Homo sapi |

Db 1280 CTGATGCGTGGAGAACTGACCGCGAGGAGCTCAAGAGATTCAAGCTGAAGCTGCTG 1339

QY 241 tcggtccgctcgcagggctacggcgcatcccgcgggcgagctgctcatggac 300

Db 1340 TCGGTGCCGTGGCGAGGCTACGGGCGATCCCGGGGCGGCTGCTGTCCATGGAC 1399

QY 301 gcttggaactcaacgacaagctggtcagcttctaccttgagactcagcgcgagctc 360

Db 1400 GCCTTGGAGCTCACGACAAGCTGTCAGCTTCTACCTGGAGACTACGGGCGGAGCTC 1459

QY 361 acgctaaagctcgcgcacatggcgctcagagatggccgggacagctcagcgcc 420

Db 1460 ACCGTAACGTGCTGGCGACATGGGCTTCAGGAGATGGCGGGGAGCTGACGGCGGC 1519

QY 421 acgcacagggtagcgcccccgttccctccaccccgctcttccctccacccacac 480

Db 1520 ACGCACAGGGTAGCGCCGCCCGCTCCCTCCACCCCGTCTTCCCTCCACCCACAC 1579

QY 481 agccttaccgcgggctctccgctttctgcttctcctaccctcctcctcctcctc 540

Db 1580 AGCGTTACCCCGGGCTCTTCGCTTTCTGTTCTCTACCTAACAAAGCTGCTC 1639

QY 541 taccggaaggaggtcccccagcttgctcaccgacacaaaggaccccgcccccagc 600

Db 1640 TACCGAAGAGGCTCCCGACGCTTGCGCTACCGACCAAGGACCCCGGCCCGCCAG 1699

QY 601 gggagggaggaagggtacactt 626

Db 1700 GGAAGGAAGGAAGGGATCACTT 1725

RESULT 2

AC009088 Homo sapiens Chromosome 16 clone RP11-388M20, WORKING DRAFT

LOCUS AC009088

DEFINITION AC009088

ACCESSION AC009088.5 GI:9256130

VERSION HTG; HTGS_PHASE2; HTGS_DRAFT.

KEYWORDS SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS DOE Joint Genome Institute.

TITLE Sequencing of Human Chromosome 16

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 233425)

AUTHORS DOE Joint Genome Institute.

TITLE Direct Submission

JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

COMMENT On Jul 18, 2000 this sequence version replaced gi:7689912.

-----Genome Center

Center: Joint Genome Institute

Center Code: JGI

Web site: <http://www.jgi.doe.gov>

Project Information

Center Project Name: 555665

Center clone name: RPCI-11_388M20

Summary Statistics

Consensus quality: 221761 bases at least Q40

Consensus quality: 229603 bases at least Q30

Consensus quality: 231140 bases at least Q20

Estimated insert size: 252540; agarose-fp estimation

Estimated insert size: 233025; sum-of-ctnigs estimation

Quality coverage: 7.01 in Q20 bases; agarose-fp estimation

Quality coverage: 7.6 in Q20 bases; sum-of-ctnigs estimation.

* NOTE: This is a 'working draft' sequence. It currently consists of 10 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.

* This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* 1 37667: contig of 37667 bp in length

* 37668 37667: gap of unknown length

* 64793 37667: contig of 27026 bp in length

* 64794 64793: gap of unknown length

* 96964 64794: contig of 32071 bp in length

* 96965 96964: gap of unknown length

* 119624 96965: contig of 22560 bp in length

* 119625 119624: gap of unknown length

* 155221 119625: contig of 35797 bp in length

* 155222 155221: gap of unknown length

* 175196 155222: contig of 19475 bp in length

* 175197 175196: gap of unknown length

* 175984 175197: contig of 788 bp in length

* 176084 175984: gap of unknown length

* 208395 176084: contig of 32311 bp in length

* 208396 208395: gap of unknown length

* 223734 208396: contig of 15239 bp in length

* 223735 223734: gap of unknown length

* 223835 223735: contig of 9591 bp in length.

Location/Qualifiers

1..233425

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="16"

/clone="RP11-388M20"

/clone_lib="RPCI human BAC library 11"

BASE COUNT 56219 a 60156 c 60134 g 56015 t 901 others

ORIGIN

Query Match 100.0%; Score 626; DB 60; Length 233425;

Best Local Similarity 100.0%; Pred. No. 1.2e-81;

Matches 626; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 agcgccctttgctggaggcaacgagcgagggcgaggtcgggagaccagagtgaggag 60

Db 10870 AGCGCCCTTTGCTGGAGAGGCAACGAGCGGGCGGGAGTCGGGAGACACAGAGTGGAGG 10929

QY 61 aagcggggagtcaggttccgcccggagcgagcttccctcctggtcgcgagctcagcg 120

Db 10930 AAGCGGGGAGTCCAGGTTCCGCCCGGAGCGGACTTCTCTCTGTCGGCGGCTCAGCG 10989

QY 121 gggtagcgcgacgacgagcgaggtcctgagcagcagcgagcgagcgagcgagcgagc 180

Db 10990 GGGTAGCGGCGGCGGCGGCGGCGGATCTGGAGCCTATGGGCGCGCGCGACGCCATC 11049

QY 181 ctgagtcgctggagaaactgaccccgagagctcagagctcagagctcagagctcagctg 240

Db 11050 CTGGATCGCTGGAGAACTGACCGCGGAGGAGCTCAAGAGTTCAAGCTGAAGCTGTG 11109

QY 241 tcggtcgcgctcgcagggctacgagcgagcgagcgagcgagcgagcgagcgagcgagc 300

Db 11110 TCGGTGCCGTGGCGGAGGGCTACGGGCGCATCCCGGGGCGGCGGCTGCTGTCCATGGAC 11169

QY 301 gcttggaactcaacgacaagctggtcagcttctaccttgagactcagcgcgagctc 360

Db 11170 GCCTTGGAGCTCACGACAAGCTGCTCAGCTTCTACCTGGAGAGCTTACGGCGCGAGCTC 11229

QY 361 accgctaacgctgctgcgcacatggcgctcagagatggcgagagctgagcgagcgagc 420

Db 11230 ACCGCTAACGCTGCTGCGGACATGGGCTGCAGGAGATGGCGGGGAGCTGACGGCGGCC 11289

QY 421 acgcacagggtagcgcccccgttccctccaccccgctcttccctccacccacac 480

Db 11290 ACGCACAGGGTAGCGCCGCCCGCTTCCCTCCACCCCGCTTTTCCCTCCACCCACAC 11349

QY 481 agcgcttaaccccggggctcttccgctttctctctctctctctctctctctctctctc 540

Db 241 TCACGACAAAGCTGGTTCAGCTTCTACCTGGAGACCTACGCGCGGAGCTCACCGGCTAACG 300
 QY 371 tctctcgacatggcctcagagatgctcggcgagctcagcgccagccacccagg 430
 Db 301 TGCTCGGACATGGGCTCGAGAGATGGCGGCGAGCTGCAGGCGGCACGCACCAAG 360
 QY 431 g 431
 Db 361 G 361

RESULT 5
 AF255794 AF255794 713 bp mRNA PRI 02-FEB-2001
 LOCUS Homo sapiens target of methylation-induced silencing-1 (TMS1) mRNA,
 DEFINITION complete cds, alternatively spliced.
 ACCESSION AF255794
 VERSION AF255794.1 GI:9802275
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 713)
 AUTHORS McConnell,B.B. and Vertino,P.M.
 TITLE Activation of a caspase-9-mediated apoptotic pathway by subcellular
 JOURNAL redistribution of the novel caspase recruitment domain protein TMS1
 MEDLINE Cancer Res. 60 (22), 6243-6247 (2000)
 PUBMED 20552140
 REFERENCE 2 (bases 1 to 713)
 AUTHORS Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.
 and Vertino,P.M.
 TITLE TMS1, a novel proapoptotic caspase recruitment domain protein, is a
 target of methylation-induced gene silencing in human breast
 JOURNAL cancers
 MEDLINE Cancer Res. 60 (22), 6236-6242 (2000)
 PUBMED 20552139
 REFERENCE 3 (bases 1 to 713)
 AUTHORS Vertino,P.M.
 TITLE Direct Submission
 JOURNAL Submitted (13-APR-2000) Radiation Oncology, Emory University School
 of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
 FEATURES
 Location/Qualifiers
 1..713
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16p11.2-p12.1"
 /cell_line="Hs578Bst"
 /tissue_type="normal breast epithelium"
 1..713
 /gene="TMS1"
 75..605
 /gene="TMS1"
 /function="involved in apoptosis"
 /codon_start=1
 /product="target of methylation-induced silencing-1"
 /protein_id="AAF99665.1"
 /db_xref="GI:9802276"
 /translation="MGRARDALDLENLTAEELKRFKLLKLVPLREGYGRIPRGAL
 LMDALDLDKLVSYFLETYGAEILTANVLRDMLQEMAGLOAATHQGLHFIHQHRAA
 LIARVTNVEMLDLYGKVLTDQYQAVRAEPTNPSKMRKLFSTFPANNWTKDLLLQ
 ALRESOSLYVEDLERS"
 ALRESOSLYVEDLERS"
 682..687
 /gene="TMS1"

polyA_signal 148 a 224 c 223 g 118 t
 BASE COUNT 148 a 224 c 223 g 118 t
 ORIGIN
 Query Match 54.88; Score 342.8; DB 89; Length 713;
 Best Local Similarity 98.08; Pred. No. 4.7e-40;
 Query Match 54.88; Score 342.8; DB 89; Length 713;
 Best Local Similarity 98.08; Pred. No. 4.7e-40;

Matches 347; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
 QY 83 ccccgagccgacttcctcctggtcgagcgtcagcgggtgagcggcgagccagccg 142
 Db 1 CCAGCGCTCGACTTCTCTCTGTTCTGGCGGCTGCAGCGGGGTGACGGGGCGACGGGCGG 60
 QY 143 gggatcctggaccatggggcgcgcgccgacgccatcctcgatgctgctgagaaacctga 202
 Db 61 GGGATCTGTGAGCCATGGGGCGCGCGACGCCATCTGTGATCGCTGGAGAACCTGA 120
 QY 203 ccggcgagcgtcaagaagttcaagctgaagctgctgctcggtcgcgctgagggct 262
 Db 121 CCGCCGAGGAGCTCAGAAGTTCAAGCTGAAGCTGCTGTGGTGGCTGCGCGAGGGCT 180
 QY 263 acggcgcatcccgcgggcgcgctgctgctcatgacgccttgacccaccacaaagc 322
 Db 181 ACGGCGCATCCCGCGGGCGCGCTGGTGTCCATGGAGCGCTTGACCTACCCACAAAGC 240
 QY 323 tggtcagcttctacctggagacctacgacgagctcacccgctcacctgctgctgagcaca 382
 Db 241 TGGTCAGCTTCTACCTGGAGACCTACGGCGCGAGCTACCGCTAACGTGCTGCGCGACA 300
 QY 383 tgggctcgagagatgcccggcgagctgagcgccgagcgccgacgagggtagc 436
 Db 301 TGGGCTGCAGGAGATGGCGGGCAGCTGCAGGCGCGCCACGACGAGGGCCTGC 354
 RESULT 6
 AF184073 AF184073 770 bp mRNA PRI 02-FEB-2001
 LOCUS Homo sapiens target of methylation-induced silencing 1 (TMS1) mRNA,
 DEFINITION complete cds.
 ACCESSION AF184073
 VERSION AF184073.1 GI:9863863
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 770)
 AUTHORS McConnell,B.B. and Vertino,P.M.
 TITLE Activation of a caspase-9-mediated apoptotic pathway by subcellular
 JOURNAL redistribution of the novel caspase recruitment domain protein TMS1
 MEDLINE Cancer Res. 60 (22), 6243-6247 (2000)
 PUBMED 20552140
 REFERENCE 2 (bases 1 to 770)
 AUTHORS Conway,K.E., McConnell,B.B., Bowring,C.E., Donald,C.D., Warren,S.T.
 and Vertino,P.M.
 TITLE TMS1, a novel proapoptotic caspase recruitment domain protein, is a
 target of methylation-induced gene silencing in human breast
 JOURNAL cancers
 MEDLINE Cancer Res. 60 (22), 6236-6242 (2000)
 PUBMED 20552139
 REFERENCE 3 (bases 1 to 770)
 AUTHORS Vertino,P.M.
 TITLE Direct Submission
 JOURNAL Submitted (09-SEP-1999) Radiation Oncology, Emory University School
 of Medicine, 145 Edgewood Avenue, SE, Atlanta, GA 30335, USA
 FEATURES
 Location/Qualifiers
 1..770
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="159538; IMAGE:1990931"
 /clone_lib="Soares breast 3NBHbst; NCI_CGAP_Ut3 uterine
 tumor"
 1..770
 /gene="TMS1"
 75..662
 /gene="TMS1"
 /function="involved in apoptosis"
 /note="target of methylation-mediated gene silencing in
 gene
 CDS


```

CDS
177. .626
/note="unnamed protein product"
/codon_start=1
/protein_id="BA91012.1"
/db_xref="GI:7020147"
/translation="MDALDLDKLVSYLETYGAELTANVLDMGLQEMAGLOLAATH
VGAAPAGIOAPPSAARPGLHFIDQHRALARTNVNELLDALYKVLTDQYOA
VRAPPTNPSKMKLFSFTPAWNWTKDLLLQALRESQSYLVEDLERS"
BASE COUNT 161 a 231 c 230 g 118 t
ORIGIN
Query Match 48.1%; Score 301; DB 89; Length 740;
Best Local Similarity 99.4%; Pred. No. 4.1e-34;
Matches 313; Conservative 0; Mismatches 0; Indels 2; Gaps 1;

Qy 117 agcggggtgagcggcgagcgccgggatactctggagccatggggcgcgcgagcgc 176
|||||
Db 1 AGCGGGGTGAGCGCGGCGAGCGCGGGATCTGAGCCATGGG--GC GCGCGCAGCG 58

Qy 177 catctggatgcgtgggaaacctgaccgcgagagctcaagaagttcaagctgaagct 236
|||||
Db 59 CATCTGGATGCGCTGGAGACCTGACCGCGGAGAGCTCAAGAAGTTCAGCTGAAGCT 118

Qy 237 gctgtcggtgcgtgcgaggggctacgggagcattcccgggcgcgctgctgtccat 296
|||||
Db 119 GCTGTGCGTGCCTGCGCGAGGGCTACGGGGCGCATCCCGGGGCGCGTGTGTCCAT 178

Qy 297 ggaagcgttgacctcaaccgaaagctggctacgttctacctggagacctacggcgccga 356
|||||
Db 179 GGACGCTTGGACCTCACCGACAAAGCTGGTCAAGCTTCTACCTGGAGACCTACGCGCGCA 238

Qy 357 gctcaccgctaagctgcgcacatggcgctgagagatggcgagagatggcgagctcagcg 416
|||||
Db 239 GCTACCGCTAACGTGCTGCCGACATGGGCTTGCAGGAGATGGCGGCGAGCTGCAGGC 298

Qy 417 ggcacgacccaggg 431
|||||
Db 299 GGCACGACCCAGG 313

RESULT 9
AX082248 AX082248 585 bp DNA PAT 27-FEB-2001
LOCUS
DEFINITION Sequence 50 from Patent WO0100826.
ACCESSION AX082248
VERSION AX082248.1 GI:13171009
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Bertin,J.
TITLE Novel molecules of the card-related protein family and uses thereof
JOURNAL Millennium Pharmaceuticals, Inc. (US)
FEATURES
Source
1..585
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 111 a 188 c 195 g 91 t
ORIGIN
Query Match 43.9%; Score 275; DB 10; Length 585;
Best Local Similarity 100.0%; Pred. No. 2.3e-30;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy. 157 atggggcgcgcgagcgcctctctgagatgcgtggagaaacctgacccgcgagagctc 216
|||||
Db 1 ATGGGGCGCGCGAGCGCATCTCTGATGCGTGGAGAACCTGACCGCGGAGAGCTC 60

Query Match 43.9%; Score 275; DB 89; Length 712;
Best Local Similarity 100.0%; Pred. No. 2.1e-30;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 157 atggggcgcgcgagcgcctctctgagatgcgtggagaaacctgacccgcgagagctc 216
|||||
Db 1 ATGGGGCGCGCGAGCGCATCTCTGATGCGTGGAGAACCTGACCGCGGAGAGCTC 60

Qy 217 aagaagttcaagctgaagctgctgcgtgcgtgcgagggctacggcgcatcccg 276
|||||
Db 61 AAGAAGTTCAAGCTGAAGCTGCTGCTGCGTGCCTGCGGAGGGCTACGGGCGCATCCG 120

Qy 277 cggggcgcgctgctgtccatgagcgccttggaacctcaccgacaaagctgctcagcttac 336
|||||
Db 121 CGGGCGCGCTGCTGTCCATGACGCCCTTGACCTACCGACAAGCTGTGTGAGCTTCTAC 180

Qy 337 ctgagagacctacggcgcgagctcaccgctaaagctgctgcgagacatgggctgcaggag 396
|||||
Db 181 CTGAGAGACCTACGGCGCGGAGCTACCGCTAACGCTGCTGCGGACATGGGCTGCGAGG 240

Qy 397 atggcgcgagcgcctcagcgcgccacacacagg 431
|||||
Db 241 ATGGCGGCGAGCTGCGAGCGGCCACGACCCAGG 275

RESULT 10
AF310103 AF310103 712 bp mRNA PRI 06-NOV-2000
LOCUS
DEFINITION Homo sapiens PYCARD mRNA, complete cds.
ACCESSION AF310103
VERSION AF310103.1 GI:11096298
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Martinon,F., Hofmann,K. and Tschopp,J.
TITLE Pycard a PYD and CARD containing molecule
JOURNAL Unpublished
2. (bases 1 to 712)
REFERENCE
AUTHORS Martinon,F., Hofmann,K. and Tschopp,J.
TITLE Direct Submission
JOURNAL Submitted (28-SEP-2000) Institute of Biochemistry, University of
Lausanne, Ch des Boveresses 155, Epalinges 1066, Switzerland
FEATURES
Source
1..712
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS
1..588
/note="contains PYD and CARD domains"
/codon_start=1
/product="PYCARD"
/protein_id="AAG30286.1"
/db_xref="GI:11096298"
/translation="MGRADAILDALENLTABELKFKLKLLSVPLRGYGRIPRGL
LSMDALDLDKLVSYLETYGAELTANVLDMGLQEMAGLOLAATHGSGAAGIOA
PQSAAKPGLHFIDQHRALARTNVNELLDALYKVLTDQYQAVRAEPTNPSKMR
KLESTPAWNWTKDLLLQALRESQSYLVEDLERS"
BASE COUNT 166 a 221 c 210 g 115 t
ORIGIN

```


[illegible]

LOCUS HUMZD54H05 556 bp mRNA PRI 29-AUG-1998
 DEFINITION Homo sapiens full length insert cDNA clone ZD54H05.
 ACCESSION AF086332
 VERSION AF086332.1 GI:3483677
 KEYWORDS FLI.CDNA.
 SOURCE human
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 556)
 Woessner,J., Tan,F., Marra,M., Kucaba,T., Yandell,M., Martin,J.,
 Math.G., Bowles,L., Wylie,T., Bowers,Y., Steptoe,M., Theising,B.,
 Geisel,S., Allen,M., Underwood,K., Chappell,J., Person,B.,
 Gibbons,M., Harvey,N., Pape,D., Chamberlain,A., Morales,R.,
 Schurk,R., Ritter,E., Kohn,S., Swaller,T., Behymer,K., Hillier,L.,
 Wilson,R. and Waterston,R.
 Full Clone Sequencing of the Longest Available Member from Each
 Unigene Cluster
 Unpublished
 2 (bases 1 to 556)
 Waterston,R.
 Direct Submission
 Submitted (24-AUG-1998) Department of Genetics, Washington
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 SUBMITTED BY:
 Genome Sequencing Center
 Department of Genetics
 Washington University
 St. Louis MO 63108, USA
 http://genome.wustl.edu/gsc
 mailto:est@watson.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PHRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

The location of this clone is unknown.

FEATURES
 source
 1..556
 /Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /Clone_lib="soares_fetal_heart_NDH19W"
 /Clone="IMAGE:344505"

BASE COUNT 124 a 168 c 177 g 87 t
 ORIGIN

Query Match 27.9%; Score 174.4; DB 97; Length 556;
 Best Local Similarity 77.8%; Pred. No. 4.9e-16;
 Matches 211; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
 QY 152 gagccatggcgcgccgacgcatcttgatgcgtggagaaacctgacccgaggg 211
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 61 GAGCCATGGGAACGAAGCGCGAGGCCATCTCTGAAGGTGCTGGAGAACCTGACACCGGAGG 120
 QY 212 agctcaagaagtcaagctgaagctgctgcgtgcgctgcgagggctacggcgca 271
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 121 AGCTCAAGAAGTCAAGATGAAGCTGGGACGGTGCCGCTGCCGAGGGGCTTGAGCGCA 180
 QY 272 tcccgggggcgcgctgctgtccatggacgcttgacctacccagacgctggtcagct 331
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 181 TCCCGGGGGCGCGCTGGGACGCTAGATATCGTGGACCTCACCGCAAGCTGGTGGCT 240
 QY 332 tctactgagacacgagcgcgagctcacgctaacgctgctgcgacatgggctgc 391
 |||| ||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 241 CCTACTACGAGGACTACGACCCGAGCTCGTCTGCGCCGCTGCTGGCGACATGGCGCATGT 300
 QY 392 aggagatggccgggcagctgcagcgccacg 423
 |||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 301 TGGAGGAGGCCGCGACGGCTGCGAGGGGCTGCG 332

Search completed: August 30, 2001, 22:30:38
 Job time: 6521 sec
